

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/29059 A2

(51) International Patent Classification⁷: C12N 15/12,
C07K 14/705

06437 (US). CATE, Richard, L. [US/US]; 40B Nichols
Road, Cohasset, MA 02025 (US). SAH, Dinah, W., Y.
[US/US]; 4 Longfellow Place, Apt. 2608, Boston, MA
02114 (US).

(21) International Application Number: PCT/US01/31488

(22) International Filing Date: 6 October 2001 (06.10.2001)

(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251
Avenue of the Americas, New York, NY 10020 (US).

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
60/238,361 6 October 2000 (06.10.2000) US

(71) Applicants (*for all designated States except US*): YALE
UNIVERSITY [US/US]; 541 College Street, New Haven,
CT 06520 (US). BIOGEN, INC. [US/US]; 14 Cambridge
Center, Cambridge, MA 02142 (US).

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): STRITTMATTER,
Stephen, M. [US/US]; 96 Tulip Tree Drive, Guilford, CT

[Continued on next page]

(54) Title: NOGO RECEPTOR HOMOLOGS

	1		50
Human NOGO-R1	-----MKRASA	GGSRLLAWVL	WLQANQVAAP CPGACVCYNE
Murine NOGO-R1	-----MKRASS	GGSRLLAWVL	WLQANRVATP CPGACVCYNE
Murine NOGO-R3	MSWQSGITVT	QSPVQAAQVS	GCCVELLLLL LAGELPLGGG CPEDVCVCYPA
Human NOGO-R3	-----	-----	-----
Human NOGO-R2	-----MLPGLRRLIQ	APASACLILM	LIALPLAAPS CPMCTCYSS
Consensus	-----	-----	CP--C-CY--
	51	LRR NT	LRR 1
Human NOGO-R1	PKVITSCPOQ	GLQAVPVGIP	AASQRIFLHG NRISHVPAAS FRACNLTIL
Murine NOGO-R1	PKVITSCPOQ	GLQAVPTGIP	ASSQRIFLHG NRISHVPAAS FQSCRNLTL
Murine NOGO-R3	P.MTVSCQAH	NFAAIPEGIP	EDSERIFLQN NRITFLOQGH FSP. AMVTL
Human NOGO-R3	-----	-----	-----
Human NOGO-R2	P.PTVSCQAN	NFSSVPLSLP	PSTQRFLQN NLIRTLRPGT FGS. NLTL
Consensus	P--T-SC--	-----P--P	-----R-FL-- N-I-----F-----L
	101	LRR 2	LRR 3
Human NOGO-R1	WLHSNVLARI	DAAAFGLAL	LEQLDLSDNA QLRSDPATF HGLGRHLTLH
Murine NOGO-R1	WLHSNVLARI	DAAAFGLTL	LEQLDLSDNA QLHVVDPTTF HGLGRHLTLH
Murine NOGO-R3	WIYSNNITFI	APNTFEGFVH	LEELDLDGDNR QLRTLAPETF QGLVKLHALY
Human NOGO-R3	WIYSNNITYI	HPSTFEGFVH	LEELDLDGDNR QLRTLAPETF QGLVKLHALY
Human NOGO-R2	WLFSNNLSTI	YPGTFRHLQ	LEELDLDGDNR HLRSLEPDTF QGLERLQSLH
Consensus	W--SN----I	----P----	LE-LDL-DN--L----P-TF -GL--L--L-
	151	LRR 4	LRR 5
Human NOGO-R1	LDRCGLELQ	PGLFRGLAAL	QYLYQDNAL QALPDDTFRD LQNLTHLFLH
Murine NOGO-R1	LDRCGLELQ	PGLFRGLAAL	QYLYQDNAL QALPDNTFRD LQNLTHLFLH
Murine NOGO-R3	LYKCGLSALP	AGIFGGLHSL	QYLYQDNHI EYLQDDIFVD LQNLTHLFLH
Human NOGO-R3	LYKCGLSALP	AGVFGGLHSL	QYLYQDNHI EYLQDDIFVD LQNLTHLFLH
Human NOGO-R2	LYRCQLSSLP	GNIFRGLVSL	QYLYQENSL LHLQDDLPAD LQNLTHLFLH
Consensus	L--C-L--L-	---P-GL--L	QYLYQ--N-- --L-D--P-D L--NL-HLFLH

(57) Abstract: The invention relates generally to genes that encode proteins that inhibit axonal growth. The invention relates specifically to genes encoding NgR protein homologs in humans and mice. The invention also includes compositions and methods for modulating the expression and activity of Nogo and the NgR proteins. Specifically, the invention includes peptides, proteins and antibodies that block Nogo-mediated inhibition of axonal extension. The compositions and methods of the invention are useful in the treatment of cranial or cerebral trauma, spinal cord injury, stroke or a demyelinating disease.

BEST AVAILABLE COPY

WO 02/29059 A2



Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOGO RECEPTOR HOMOLOGS

FIELD OF THE INVENTION

5 The invention relates to neurology and molecular biology. More particularly, the invention relates to CNS neurons and axonal growth

BACKGROUND

10 Among the mechanisms through which the cells of an organism communicate with each other and obtain information and stimuli from their environment is through cell membrane receptor molecules expressed on the cell surface. Many such receptors have been identified, characterized, and sometimes classified into major receptor superfamilies based on structural motifs and signal transduction features. The receptors are a first essential link for translating an extracellular signal into a cellular
15 physiological response.

 Receptors on neurons are particularly important in the development of the nervous system during embryogenesis. The neurons form connections with target cells during development through axonal extension of the neurons toward the target cells in a receptor-mediated process. Axons and dendrites have a specialized region of their
20 distal tips known as the growth cone. Growth cones enable the neuron to sense the local environment through a receptor-mediated process and direct the movement of the axon or dendrite of the neuron toward the neuron's target cell. This process is known as elongation. Growth cones can be sensitive to several guidance cues, for example, surface adhesiveness, growth factors, neurotransmitters and electric fields. The
25 guidance of growth at the cone depends on various classes of adhesion molecules, intercellular signals, as well as factors that stimulate and inhibit growth cones.

 Interestingly, damaged neurons do not elongate in the central nervous system (CNS) following injury due to trauma or disease, whereas axons in the peripheral nervous system (PNS) regenerate readily. The fact that damaged CNS neurons fail to
30 elongate is not due to an intrinsic property of CNS axons, but rather due to the CNS environment that is not permissive for axonal elongation. Classical grafting experiments by Aguayo and colleagues (*e.g.*, Richardson *et al.*, (1980) *Nature* 284,

- 2 -

264-265) demonstrated that CNS axons can in fact elongate over substantial distances within peripheral nerve grafts, and that CNS myelin inhibits CNS axon elongation. Therefore, given the appropriate environment, CNS axons can regenerate, implying that CNS axonal injury can potentially be addressed by appropriate manipulation of the
5 CNS environment.

The absence of axon regeneration following injury can be attributed to the presence of axon growth inhibitors. These inhibitors are predominantly associated with myelin and constitute an important barrier to regeneration. Axon growth inhibitors are present in CNS-derived myelin and the plasma membrane of
10 oligodendrocytes that synthesize myelin in the CNS (Schwab *et al.*, (1993) *Annu. Rev. Neurosci.* 16, 565-595). Myelin-associated inhibitors appear to be a primary contributor to the failure of CNS axon regeneration *in vivo* after an interruption of axonal continuity, whereas other non-myelin associated axon growth inhibitors in the CNS may play a lesser role. These inhibitors block axonal regeneration following
15 neuronal injury due to trauma, stroke or viral infection.

Numerous myelin-derived axon growth inhibitors have been characterized (see, for review, David *et al.*, (1999) WO995394547; Bandman *et al.*, (1999) U.S. Patent No. 5,858,708; Schwab, (1996) *Neurochem. Res.* 21, 755-761). Several components of CNS white matter, NI35, NI250 (Nogo) and Myelin-associated glycoprotein
20 (MAG), which have inhibitory activity for axonal extension, have been described as well (Schwab *et al.*, (1990) WO9005191; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). In particular, Nogo is a 250 kDa myelin-associated axon growth inhibitor that was originally characterized based on the effects of the purified protein *in vitro* and monoclonal antibodies that neutralize the protein's activity (Schwab (1990) *Exp.*
25 *Neurol.* 109, 2-5). The Nogo cDNA was first identified through random analysis of brain cDNA and had no suggested function (Nagase *et al.*, (1998) *DNA Res.* 5, 355-364). The identification of this Nogo cDNA as the cDNA encoding the 250 kDa myelin-associated axon growth inhibitor was discovered only recently (GrandPre *et al.*, (2000) *Nature* 403, 439-444; Chen *et al.*, (2000) *Nature* 403, 434-439; Prinjha *et al.*,
30 (2000) *Nature* 403, 383-384).

Importantly, Nogo has been shown to be the primary component of CNS myelin responsible for inhibiting axonal elongation and regeneration. Nogo's selective

expression by oligodendrocytes and not by Schwann cells (the cells that myelinate P.S. axons) is consistent with the inhibitory effects of CNS myelin, in contrast to P.S. myelin (GrandPre *et al.*, (2000) *Nature* 403, 434-439). In culture, Nogo inhibits axonal elongation and causes growth cone collapse (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272, 19283-19293). Antibodies (*e.g.*, IN-1) against Nogo have been shown to block most of the inhibitory action of CNS myelin on neurite growth *in vitro* (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272:19283-19293). These experiments indicate that Nogo is the main component of CNS myelin responsible for inhibition of axonal elongation in culture. Furthermore, *in vivo*, the IN-1 antibody has been shown to enhance axonal regeneration after spinal cord injury, resulting in recovery of behaviors such as contact placing and stride length (Schnell and Schwab (1990) *Nature* 343, 269-272; Bregman *et al.*, (1995) *Nature* 378, 498-501). Thus, there is substantial evidence that Nogo is a disease-relevant molecular target. Agents that interfere with the binding of Nogo to its receptor would be expected to improve axonal regeneration in clinical states in which axons have been damaged, and improve patient outcome.

Modulation of Nogo has been described as a means for treatment of regeneration for neurons damaged by trauma, infarction and degenerative disorders of the CNS (Schwab *et al.*, (1994) WO9417831; Tatagiba *et al.*, (1997) *Neurosurgery* 40, 541-546) as well as malignant tumors in the CNS such as glioblastoma (Schwab *et al.*, (1993) U.S. Patent No. 5,250,414; Schwab *et al.*, (2000) U.S. Patent No. 6,025,333).

Antibodies which recognize Nogo have been suggested to be useful in the diagnosis and treatment of nerve damage resulting from trauma, infarction and degenerative disorders of the CNS (Schnell & Schwab, (1990) *Nature* 343, 269-272; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). For CNS axons, there is a correlation between the presence of myelin and the inhibition of axon regeneration over long distances (Savio and Schwab (1990) *Proc. Natl. Acad. Sci.* 87, 4130-4133; Keirstead *et al.*, (1992) *Proc. Natl. Acad. Sci.* 89, 11664-11668). After Nogo is blocked by antibodies, neurons can again extend across lesions caused by nerve damage (Schnell and Schwab (1990) *Nature* 343, 269-272).

- 4 -

SUMMARY OF THE INVENTION

Genes encoding homologs (NgR2 and NgR3) of a Nogo receptor (NgR1) in mice and humans have been discovered. Various domains in the polypeptides encoded by the NgR2 and NgR3 genes have been identified and compared to domains in mouse
 5 and human NgR1 polypeptides. This comparison has led to identification of a consensus sequence (NgR consensus sequence) that characterizes a family of proteins (NgR family). Based on these and other discoveries, the invention features molecules and methods for modulating axonal growth in CNS neurons.

The invention provides a polypeptide that contains a polypeptide containing a
 10 tryptophan rich LRRCT domain consisting of the amino acid sequence:

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

15

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any protein amino acid or a gap, and the polypeptide does not include amino acid sequence from residue 260 to 309 of SEQ ID NO: 5
 20 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

Preferably, X17 and X23 are (independently) arginine or lysine. In some embodiments, the amino acid sequence of the LRRCT domain is residues 261-310 of SEQ ID NO:2, or residues 261-310 of SEQ ID NO: 2 with up to 10 conservative
 25 amino acid substitutions. In some embodiments, the polypeptide contains the following NTLRRCT amino acid sequence:

C P X₁ X₂ C X₃ C Y X₄ X₅ P X₆ X₇ T X₈ S C X₉ X₁₀ X₁₁ X₁₂ X₁₃ X₁₄ X₁₅ X₁₆ P
 X₁₇ X₁₈ X₁₉ P X₂₀ X₂₁ X₂₂ X₂₃ R X₂₄ F L X₂₅ X₂₆ N X₂₇ I X₂₈ X₂₉ X₃₀ X₃₁ X₃₂ X₃₃
 X₃₄ F X₃₅ X₃₆ X₃₇ X₃₈ X₃₉ X₄₀ X₄₁ X₄₂ L W X₄₃ X₄₄ S N X₄₅ X₄₆ X₄₇ X₄₈ I X₄₉
 30 X₅₀ X₅₁ X₅₂ F X₅₃ X₅₄ X₅₅ X₅₆ X₅₇ L E X₅₈ L D L X₅₉ D N X₆₀ X₆₁ L X₆₂ X₆₃ X₆₄
 X₆₅ P X₆₆ T F X₆₇ G L X₆₈ X₆₉ L X₇₀ X₇₁ L X₇₂ L X₇₃ X₇₄ C X₇₅ L X₇₆ X₇₇ L X₇₈
 X₇₉ X₈₀ X₈₁ F X₈₂ G L X₈₃ X₈₄ L Q Y L Y L Q X₈₅ N X₈₆ X₈₇ X₈₈ X₈₉ L X₉₀ D

- 5 -

X₉₁ X₉₂ F X₉₃ D L X₉₄ N L X₉₅ H L F L H G N X₉₆ X₉₇ X₉₈ X₉₉ X₁₀₀ X₁₀₁ X₁₀₂
 X₁₀₃ X₁₀₄ F R G L X₁₀₅ X₁₀₆ L D R L L L H X₁₀₇ N X₁₀₈ X₁₀₉ X₁₁₀ X₁₁₁ V H X₁₁₂
 X₁₁₃ A F X₁₁₄ X₁₁₅ L X₁₁₆ R L X₁₁₇ X₁₁₈ L X₁₁₉ L F X₁₂₀ N X₁₂₁ L X₁₂₂ X₁₂₃ L
 X₁₂₄ X₁₂₅ X₁₂₆ X₁₂₇ L X₁₂₈ X₁₂₉ L X₁₃₀ X₁₃₁ L X₁₃₂ X₁₃₃ L R L N X₁₃₄ N X₁₃₅ W
 5 X₁₃₆ C X₁₃₇ C R X₁₃₈ R X₁₃₉ L W X₁₄₀ W X₁₄₁ X₁₄₂ X₁₄₃ X₁₄₄ R X₁₄₅ S S S X₁₄₆
 V X₁₄₇ C X₁₄₈ X₁₄₉ P X₁₅₀ X₁₅₁ X₁₅₂ X₁₅₃ X₁₅₄ X₁₅₅ D L X₁₅₆ X₁₅₇ L X₁₅₈ X₁₅₉ X₁₆₀
 D X₁₆₁ X₁₆₂ X₁₆₃ C [SEQ ID NO:18]

wherein X is any amino acid residue or a gap and wherein the polypeptide is not the
 10 polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1). For
 example, X₆, X₃₇ and X₃₈ may represent a gap. Specific examples of polypeptides of
 the invention are SEQ ID NO: 2 (human NgR2), SEQ ID NO: 4 (mouse NgR3), and
 SEQ ID NO: 14 (human NgR3). In some embodiments, the polypeptide contains: (a)
 a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial
 15 CTS domain, if present, consists of no more than the first 39 amino acids of the CTS
 domain. While the polypeptide may contain a functional GPI domain, a functional GPI
 domain may be absent, e.g., when a soluble polypeptide is desired. A polypeptide of
 the invention optionally includes an amino acid sequence of a heterologous
 polypeptide, e.g., an Fc portion of an antibody.

20 The invention also provides a nucleic acid encoding an above-described
 polypeptide; a vector containing the nucleic acid, which nucleic acid may be operably
 linked to an expression control sequence; and a transformed host cell containing the
 vector. A method of producing a polypeptide of the invention is also provided. The
 method includes introducing a nucleic acid encoding the above-described polypeptide
 25 into a host cell, culturing the cell under conditions suitable for expression of the
 polypeptide, and recovering the polypeptide.

The invention also provides an antisense molecule whose nucleotide sequence
 is complementary to a nucleotide sequence encoding a polypeptide selected from the
 group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a
 30 polypeptide consisting of residues 256-396 of SEQ ID NO:14 and a polypeptide
 consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to

- 6 -

100 nucleotides in length, e.g., about 20, 30, 40, 50, 60, 70, 80 or 90 nucleotides. The invention also provides a nucleic acid encoding such an antisense molecule.

The invention also provides an antibody that binds to an above-described polypeptide. Polypeptides or antibodies of the invention can be formulated into
5 pharmaceutical compositions containing the polypeptide or antibody and a pharmaceutically acceptable carrier.

The invention also provides a method for decreasing inhibition of axonal growth of a CNS neuron. The method includes the step of contacting the neuron with an effective amount of a polypeptide or antibody of the invention.

10 The invention also provides a method for treating a central nervous system disease, disorder or injury. The method includes administering to a mammal, e.g., a human, an effective amount of a polypeptide or antibody of the invention. Exemplary diseases, disorders and injuries that may be treated using molecules and methods of the invention include, but are not limited to, cerebral injury, spinal cord injury, stroke,
15 demyelinating diseases, e.g., multiple sclerosis, monophasic demyelination, encephalomyelitis, multifocal leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease, Spongy degeneration, Alexander's disease, Canavan's disease, metachromatic leukodystrophy and Krabbe's disease.

The invention also provides a method for identifying a molecule that binds a
20 polypeptide of the invention. The method includes the steps of: (a) providing a polypeptide of the invention; (b) contacting the polypeptide with the candidate molecule; and (c) detecting binding of the candidate molecule to the polypeptide.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the
25 invention belongs. In case of conflict, the present application, including definitions, will control. All publications, patent and other references mentioned herein are incorporated by reference.

The materials, methods and examples presented below are illustrative only, and not intended to be limiting. Other features and advantages of the invention will be
30 apparent from the detail description and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

- 7 -

Fig. 1A-1B shows an alignment of NgR2 (SEQ ID NO:2) and NgR3 (SEQ ID NO:4) with the known NgR, NgR1 (SEQ ID NO:5) and the Consensus Sequence (SEQ ID NO:6).

5

Fig. 2. mNgR3 does not bind hNogoA(1055-1120). COS-7 cells were transfected with vectors encoding myc-NgR1 or myc-NgR3, fixed, and stained with anti-myc antibodies or AP-hNogoA(1055-1120).

10 Fig.3. An alignment of the amino acid sequences of human NgR1, murine NgR1, murine NgR3, human NgR3 and human NgR2. Numbering begins with amino acid #1 of murine NgR3. The consensus sequence is listed below. The LRR NT domain is indicated by a shaded box; domains LLR 1, LLR 3, LLR 5, and LLR 7 are indicated by open boxes; LLR 2, LLR 4, LLR 6 and LLR 8 are indicated by shaded
15 boxes; and the LLR CT domain is indicated by a shaded box. Amino acids in bold in LLR 8 indicate a conserved glycosylation sites. A dot indicates conserved cysteine residue in LLR4. Box at C terminus indicates putative GPI signals.

20

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) encoding NgR homologs, referred to herein as NgR. Unless indicated otherwise, as used herein,
25 the abbreviation in lower case (NgR) refers to a gene, cDNA, RNA or nucleic acid sequence, whereas the upper case version (NgR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence. Specific proteins are designated by number, *e.g.*, "NgR2" is a human NgR homolog, "NgR3" is a murine-derived NgR homolog, and "NgR1" is the known NgR identified by Dr. Stephen Strittmatter.
30 Known NgRs are herein referred to as "NgRs." DNA polynucleotides of the invention include genomic DNA, cDNA and DNA that has been chemically synthesized in whole or in part.

Standard reference works setting forth the general principles of recombinant DNA technology known to those of skill in the art include Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York (1998); Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d Ed., Cold
5 Spring Harbor Laboratory Press, Plainview, New York (1989); Kaufman *et al.*, Eds., HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton (1995); McPherson, Ed., DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford (1991).

As used herein, the term "axon" refers to a long cellular protrusion from a
10 neuron, whereby action potentials are conducted, either to or from the cell body.

As used herein, the term "axonal growth" refers to an extension of the long process or axon, originating at the cell body and proceeded by the growth cone.

As used herein, the term "central nervous system disorder" refers to any pathological state associated with abnormal function of the central nervous system
15 (CNS). The term includes, but is not limited to, altered CNS function resulting from physical trauma to cerebral tissue, viral infection, autoimmune mechanisms and genetic mutation.

As used herein, the term "demyelinating disease" refers to a pathological disorder characterized by the degradation of the myelin sheath of the oligodendrocyte
20 cell membrane.

As used herein, the term "growth cone" refers to a specialized region at the tip of a growing neurite that is responsible for sensing the local environment and moving the axon toward its appropriate synaptic target cell.

As used herein, the term "growth cone movement" refers to the extension or
25 collapse of the growth cone toward a neuron's target cell.

As used herein, the term "neurite" refers to a process growing out of a neuron. As it is sometimes difficult to distinguish a dendrite from an axon in culture, the term "neurite" is used for both.

As used herein, the term "oligodendrocyte" refers to a neuroglial cell of the
30 CNS whose function is to myelinate CNS axons.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods.

- 9 -

"Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means. By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of
5 proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein. The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in
10 addition to all or part of that region. Examples of NgR protein domains include, but are not limited to, the signal peptide, extracellular (*i.e.*, N-terminal) domain, and leucine-rich repeat domains.

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, *i.e.*,
15 having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event. Such activities may be measured by assays such as competitive inhibition of NgR1 binding to
20 Nogo assays wherein, for example, unlabeled, soluble NgR2 is added to an assay system in increasing concentrations to inhibit the binding of Nogo to NgR1 expressed on the surface of CHO cells. As another example, one may assess the ability of neurons to extend across lesions caused by nerve damage (as in Schnell and Schwab (1990) *Nature* 343, 269-272) following inhibition of Nogo by various forms of NgR2
25 and/or NgR3 as a biological indicator of NgR function.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)₂, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies, anti-idiotypic antibodies, anti-anti-idiotypic antibodies, and
30 humanized antibodies.

As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or

- 10 -

combinations thereof. Binding includes ionic, non-ionic, hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or
5 due to the effect of another protein or compound but instead are without other substantial chemical intermediates.

As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecules, peptides, proteins, sugars, nucleotides or nucleic acids, and such compound can be natural or synthetic.

10 As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers,
15 salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the NgR or fragment thereof.

As used herein, the phrase "homologous nucleotide sequence," or "homologous
20 amino acid sequence," or variations thereof, refers to sequences characterized by an identity at the nucleotide level, or a homology at the amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA.

25 Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does
30 not, however, include the nucleotide sequence encoding NgR1. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. A

- 11 -

homologous amino acid sequence does not, however, include the amino acid sequence encoding other known NgRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default
5 settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that is substantially free of nucleic acids encoding other proteins with which it is associated in nature, i.e., a nucleic acid that has been removed
10 from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*,
15 sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NgR nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the
20 nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

As used herein, the term "heterologous" refers to a nucleotide or amino acid
25 sequence that is a different, or non-corresponding sequence, or a sequence derived from a different species. For example, a mouse NgR nucleotide or amino acid sequence is heterologous to a human NgR nucleotide or amino acid sequence, and a human NgR nucleic or amino acid sequence is heterologous to a human immunoglobulin nucleotide or amino acid sequence.

30 As used herein, a "soluble NgR polypeptide" is a NgR polypeptide that does not anchor itself in a membrane. Such soluble polypeptides include, for example, NgR2 and NgR3 polypeptides that lack a sufficient portion of their GPI anchor signal

- 12 -

to anchor the polypeptide or are modified such that the GPI anchor signal is not adequate to result in replacement of the peptide with a GPI anchor. In preferred embodiments, up to 5, 10, 20 or 25 amino acids are removed from the C-terminus of NgR2 or NgR3 to make the respective proteins soluble. As used herein soluble NgR polypeptides include full-length or truncated (*e.g.*, with internal deletions) NgR.

Soluble NgR polypeptides may include the entire NgR protein up to the putative GPI signal sequence (*e.g.*, amino acid 1 to about amino acid 395 of NgR2, and from amino acid 1 to about amino acid 438 of NgR3). In other embodiments, the signal peptide of the proteins may be removed or truncated (*e.g.*, all or part of the signal sequence of NgR2, which spans amino acid 1 to about amino acid 30 of SEQ ID NO:2, may be removed; all or part of the signal sequence of NgR3, which spans amino acid 1 to about amino acid 40 of SEQ ID NO:4, may be removed). In some embodiments, the mature NgR2 (SEQ ID NO:8) and the mature NgR3 (SEQ ID NO:9) are used.

Soluble NgR polypeptides include at least one of the putative ligand-binding portions of NgR, including the first cysteine-rich region (SEQ ID NO:10, the leucine repeat region (SEQ ID NO:12) and the second cysteine-rich region (SEQ ID NO:11). In some embodiments, soluble NgR polypeptides consist of amino acid 1 through about amino acid 395 of SEQ ID NO:2, or amino acid 1 through about amino acid 438 of SEQ ID NO:4.

In other embodiments, the soluble NgR polypeptides are fusion proteins that contain amino acids 30 through about amino acid 395 of mature NgR2 or amino acid 40 through about amino acid 438 of NgR3, the C-terminal 10 amino acids of a human IgG 1 hinge region containing the two cysteine residues thought to participate in interchain disulfide bonding, and the CH2 and CH3 regions of a human IgG1 heavy chain constant domain. This type of recombinant protein is designed to modulate inhibition of axonal elongation through inhibition of the Nogo ligand binding to NgR1, or by inhibiting the ligand of the NgR from interacting with cell surface NgR. The NgR portion of the fusion binds to the Nogo ligand and the IgG1 portion binds to the FcγRI (macrophage) and FcγIII (NK cells and neutrophils) receptors.

The production of the soluble polypeptides useful in this invention may be achieved by a variety of methods known in the art. For example, the polypeptides may

- 13 -

be derived from intact transmembrane NgR molecules by proteolysis using specific endopeptidases in combination with exopeptidases, Edman degradation, or both. The intact NgR molecule, in turn, may be purified from its natural source using conventional methods. Alternatively, the intact NgR may be produced by known
5 recombinant DNA techniques using cDNAs, expression vectors and well-known techniques for recombinant gene expression.

Preferably, the soluble polypeptides useful in the present invention are produced directly, thus eliminating the need for an entire NgR as a starting material. This may be achieved by conventional chemical synthesis techniques or by well-known
10 recombinant DNA techniques wherein only those DNA sequences which encode the desired peptides are expressed in transformed hosts. For example, a gene which encodes the desired soluble NgR polypeptide may be synthesized by chemical means using an oligonucleotide synthesizer. Such oligonucleotides are designed based on the amino acid sequence of the desired soluble NgR polypeptide. Specific DNA sequences
15 coding for the desired peptide also can be derived from the full-length DNA sequence by isolation of specific restriction endonuclease fragments or by PCR synthesis of the specified region from cDNA.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NOs: 1, 3 or a complement of either of
20 these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NOs: 1 or 3 as a hybridization probe, NgR nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY
25 MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993).

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers
30 according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

Furthermore, oligonucleotides corresponding to NgR nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

5 As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue.

10 Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

 As used herein, the term "probe" refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides,
15 depending on use. They are used in the detection of identical, similar or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

20 The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

 The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

 The term "therapeutic effect" refers to the inhibition or activation factors
25 causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of
30 degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population

- 15 -

of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal
5 condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions, for example, include cancers such as
10 fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus and inflammation.

Abnormal differentiation conditions include, for example, neurodegenerative disorders, slow wound healing rates and slow tissue grafting healing rates.

Abnormal cell signaling conditions include, for example, psychiatric disorders
15 involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature
20 cell death.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell
25 culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

30 The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then

- 16 -

be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable
5 presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1–2-fold, and preferably more, compared to the basal level.

The amino acid sequences are presented in the amino to carboxy direction,
10 from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

15

Nucleic Acids

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA
20 transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a NgR polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the
25 invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild-type genes, are naturally occurring sequences (as
30 opposed to non-naturally occurring variants arising from in vitro manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding NgR (conventionally followed by

- 17 -

second-strand synthesis of a complementary strand to provide a double-stranded DNA).

Preferred DNA sequences encoding a human NgR polypeptide is set out in SEQ ID NOs:1 and 13. A preferred DNA of the invention comprises a double
5 stranded molecule comprising the coding molecule (*i.e.*, the "coding strand") along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding NgR polypeptides, as shown in SEQ ID NO:3, which comprises murine NgR homolog,
10 NgR3.

Also preferred are nucleotide sequences that encode at least a portion of a NgR polypeptide that has at least one biological function of a NgR. More preferred are nucleotide sequences that encode a portion of NgR that encodes at least the mature NgR without the hydrophobic C-terminal GPI signal. Also preferred are nucleotide
15 sequences that encode the portion of NgR that encodes at least the ligand-binding region of NgR.

The invention further embraces other species, preferably mammalian, homologs of the human NgR DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at
20 least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with human DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the NgR sequences set forth in SEQ ID
25 NOs:1, 3 or 13, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art. Polynucleotides of the invention also permit
30 identification and isolation of polynucleotides encoding related NgR polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR).

- 18 -

Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to NgR and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of NgR. Non-human species genes encoding proteins homologous to NgR can also be identified by Southern and/or PCR analysis and are useful in animal models for NgR disorders. Knowledge of the sequence of a human NgR DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding NgR expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express NgR. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a NgR locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

The disclosure herein of a full-length polynucleotide encoding a NgR polypeptide makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide. The invention, therefore, provides fragments of NgR-encoding polynucleotides comprising at least 6, and preferably at least 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding NgR. Preferably, fragments of polynucleotides of the invention comprise sequences unique to the NgR-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding NgR (or fragments thereof).

Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic

- 19 -

DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent and enzymatic labeling.

5 Fragments of polynucleotides are particularly useful as probes for detection of full-length or fragment of NgR polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding NgR, or used to detect variations in a polynucleotide sequence encoding NgR.

The invention also embraces DNAs encoding NgR polypeptides that hybridize under moderately stringent or high stringency conditions to the noncoding strand, or
10 complement, of the polynucleotide in any of SEQ ID NOs:1 or 3.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1?6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98% or 99% homologous to each other typically remain
15 hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that
20 hybridizes under stringent conditions to the sequence of SEQ ID NOs:1 or 3 corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein). As used herein, "stringent hybridization conditions" means: 42°C in a hybridization
25 solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% (wt/vol) dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS.

Vectors

30 Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding NgR and/or to

- 20 -

express DNA which encodes NgR. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), that serve equivalent functions.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67, 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia,

- 21 -

Piscataway, N.J.) that fuse glutathione-S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69, 301-315) and pET 11d (Studier *et al.*, GENE
5 EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY:
10 METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20, 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by
15 standard DNA synthesis techniques.

In another embodiment, the NgR expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J.* 6, 229-234), pMFa (Kurjan and Herskowitz (1982) *Cell* 30, 933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54, 113-123), pYES2
20 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, NgR can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith *et al.*, (1983) *Mol. Cell. Biol.* 3,
25 2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170, 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329, 840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6, 187-195). When used in mammalian cells, the
30 expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both

- 22 -

prokaryotic and eukaryotic cells. See, e.g., Chapters 16 and 17 of Sambrook *et al.*, (Eds.) MOLECULAR CLONING: A LABORATORY MANUAL. 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is
5 capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1, 268-277), lymphoid-specific promoters (Calame and Eaton
10 (1988) *Adv. Immunol.* 43, 235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8, 729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33, 729-740; Queen and Baltimore (1983) *Cell* 33, 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86, 5473-5477), pancreas-specific promoters (Edlund *et al.*
15 (1985) *Science* 230, 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss (1990) *Science* 249, 374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3, 537-546).

20 The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense NgR mRNA. Regulatory sequences operatively linked
25 to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue-specific or cell-type-specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid,
30 phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation

of gene expression using antisense genes see Weintraub *et al.*, *Antisense RNA as a molecular tool for genetic analysis*, REVIEWS--TRENDS IN GENETICS, Vol. 1(1) 1986.

Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses and integratable DNA fragments (*i.e.*, fragments
5 integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors
10 include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen) and pYES2™ (Invitrogen).

Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding NgR is operably linked or connected to suitable control sequences
15 capable of effecting the expression of the NgR in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually
20 conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include, but are not limited to a transcriptional promoter, enhancers, an optional operator sequence to control transcription, polyadenylation signals, a
25 sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation. Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types
30 of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as

- 24 -

the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, NgR proteins, mutant forms of NgR, fusion proteins, etc.).

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the PR and PL promoters of bacteriophage lambda (THE BACTERIOPHAGE LAMBDA, Hershey, A.D. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; LAMBDA II, Hendrix, R.W. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the *trp*, *recA*, heat shock, and *lacZ* promoters of *E. coli* and the SV40 early promoter (Benoist *et al.*, (1981) *Nature* 290, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine and human metallothionein.

Additional regulatory sequences can also be included in preferred vectors. Preferred examples of suitable regulatory sequences are represented by the Shine-Dalgarno sequence of the replicase gene of the phage MS-2 and of the gene *cII* of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding NgR and result in the expression of the mature NgR protein.

Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors

- 25 -

which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and NgR DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

5 Nucleotide sequences encoding NgR may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and ligation with appropriate ligases. Techniques for such manipulation are
10 disclosed by Sambrook *et al.*, *supra* and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, (1983) *Mol. Cell. Biol.* 3:280, Cosman *et al.* (1986) *Mol. Immunol.* 23:935, Cosman *et al.*, (1984) *Nature* 312:768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

15

Host Cells and Transformed Host Cells

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded NgR
20 polypeptide. Preferably, the cell produces little or no endogenous NgR polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation,
25 nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate and mammalian cells systems.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with NgR. Host cells of the
30 invention are also useful in methods for the large-scale production of NgR polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells

- 26 -

are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of NgR DNA sequences allows for modification of cells to permit, or increase, expression of endogenous NgR. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring NgR promoter with all or part of a heterologous promoter so that the cells express NgR at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous NgR encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the NgR coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the NgR coding sequences in the cells.

The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination or "knock-out" strategies; see Capecchi, *Science* 244:1288-1292 (1989)) of animals that fail to express functional NgR or that express a variant of NgR. Such animals (especially small laboratory animals such as rats, rabbits and mice) are useful as models for studying the *in vivo* activities of NgR and modulators of NgR.

- 27 -

Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces* and *Staphylococcus*.

If a eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*, Tissue Culture, Academic Press, Kruse and Patterson, Eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast cell may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia* and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, 1988, 6, 47; BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, 1992; and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBACTTM complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA

- 28 -

(1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via
5 conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or
10 transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells
15 may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin, dihydrofolate reductase (DHFR) and methotrexate. Nucleic acid
20 encoding a selectable marker can be introduced into a host cell on the same vector as that encoding NgR or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

In a preferred embodiment, the polypeptides of the invention, including forms
25 of NgR2 and NgR3, soluble forms of NgR, chimeric NgR polypeptides, NgR/Ig fusions and fragments and variations of each of the above are expressed in Chinese Hamster Ovary (CHO) cells.

In order to introduce the DNA fragment coding for the NgR protein or polypeptide into the CHO cell to express the recombinant NgR protein or polypeptide,
30 it is necessary to construct the expression vector.

The vectors for CHO expression include, but are not limited to, pA1-11, pXT1, pRc/CMV, pRc/RSV and pcDNAINeo. The promoter is not specifically limited

- 29 -

provided it effectively promotes expression in CHO cells. Examples of suitable promoters are: SR α , SV40, LTR, CMV, and HSV-TK. Of these, CMV and SR α promoters are preferred.

5 In addition to the above-mentioned promoters, the expression vectors may contain enhancers, splicing signals, polyadenylation signals, selectable markers and an SV40 replication origin. Suitable selectable markers include, but are not limited to the dihydrofolate reductase (DHFR) gene which provides resistance to methotrexate (MTX), the ampicillin resistance gene, and the neomycin resistance gene.

10 Examples of the expression vectors each containing the DNA coding for NgR, portions, fragments and soluble constructs thereof, include the vector (such as one described above), into which the promoter is operably linked (preferably upstream) to the nucleotide sequence encoding the desired NgR construct; a polyadenylation signal downstream from the nucleotide sequence encoding the NgR construct; and, preferably, the vector includes an operable DHFR gene. Preferably, the ampicillin
15 resistant gene is also operably contained in the vector.

CHO cell lacking the DHFR gene (Urlaub, G. *et al.*, (1980) *Proc. Natl. Acad. Sci. USA* 77, 4216-4220) and CHO-K1 (*Proc. Natl. Acad. Sci. USA* 60, 1275 (1968)) are suitable for use.

20 The NgR expression vectors prepared as above are introduced into CHO cells by any known method, including, but not limited to the calcium phosphate method (Graham and van der Eb (1973) *Viol.* 52, 456-467) and electroporation (Nuemann *et al.*, (1982) *EMBO J.* 1, 841-845).

25 Transformants carrying the expression vectors are selected based on the above-mentioned selectable markers. Repeated clonal selection of the transformants using the selectable markers allows selection of stable cell lines having high expression of the NgR constructs. Increased MTX concentrations in the selection medium allows gene amplification and greater expression of the desired protein. The CHO cell containing the recombinant NgR can be produced by cultivating the CHO cells containing the NR expression vectors constitutively expressing the NgR constructs.

30 Media used in cultivating CHO cells includes DMEM medium supplemented with about 0.5 to 20% fetal calf serum, DMEM medium and RPMI1640 medium. The

- 30 -

pH of the medium is preferably about 6 to 8. Cultivation is preferably at about 30 to 40°C for about 15 to 72 hours with aeration.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NgR protein. Accordingly, the invention further provides methods for producing NgR protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding NgR has been introduced) in a suitable medium such that NgR protein is produced. In another embodiment, the method further comprises isolating NgR from the medium or the host cell.

In situations where the NgR polypeptide will be found primarily intracellularly, intracellular material (including inclusion bodies for Gram-negative bacteria) can be extracted from the host cell using any standard technique known to one of ordinary skill in the art. Such methods would encompass, by way of example and not by way of limitation, lysing the host cells to release the contents of the periplasm/cytoplasm by French press, homogenization and/or sonication followed by centrifugation.

If the NgR polypeptide has formed inclusion bodies in the cytosol, such inclusion bodies may frequently bind to the inner and/or outer cellular membranes. Upon centrifugation, the inclusion bodies will be found primarily in the pellet material. The pellet material can then be treated at pH extremes or with one or more chaotropic agents such as a detergent, guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris-carboxyethyl phosphine at acid pH to release, break apart and solubilize the inclusion bodies. Once solubilized, NgR polypeptide can be analyzed using gel electrophoresis, immunoprecipitation or the like. Various methods of isolating the NgR polypeptide would be apparent to one of ordinary skill in the art, for example, isolation may be accomplished using standard methods such as those set forth below and in Marston *et al* (1990) *Meth. Enzymol.* 182, 264-275 (incorporated by reference herein in its entirety).

If isolated NgR polypeptide is not biologically active following the isolation procedure employed, various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages, can be used to restore biological

- 31 -

activity. Methods known to one of ordinary skill in the art include adjusting the pH of the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization but usually at a lower concentration and is not necessarily the same chaotrope as used for the solubilization. It may be required to employ a reducing agent or the reducing agent plus its oxidized form in a specific ratio, to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridge(s). Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol (DTT)/dithiane DTT, 2-mercaptoethanol (bME)/dithio-b(ME). To increase the efficiency of the refolding, it may be necessary to employ a cosolvent, such as glycerol, polyethylene glycol of various molecular weights and arginine.

Transgenic Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NgR-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NgR sequences have been introduced into their genome or homologous recombinant animals in which endogenous NgR sequences have been altered. Such animals are useful for studying the function and/or activity of NgR and for identifying and/or evaluating modulators of NgR activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous NgR gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA

- 32 -

molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NgR-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by
5 microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human NgR DNA sequence of SEQ ID NOs:1 or 3 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homolog of the human NgR gene, such as a mouse NgR gene, can be isolated based on hybridization to the human NgR cDNA (described
10 further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the NgR transgene to direct expression of NgR protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly
15 animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, in MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the
20 NgR transgene in its genome and/or expression of NgR mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding NgR can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which
25 contains at least a portion of a NgR gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NgR gene. The NgR gene can be a human gene (*e.g.*, SEQ ID NOs:1 or 13), but more preferably, is a non-human homolog of a human NgR gene. For example, a mouse homolog of human NgR gene of SEQ ID NOs:1 or 13 can be used to construct a homologous
30 recombination vector suitable for altering an endogenous NgR gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous

- 33 -

recombination, the endogenous NgR gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NgR gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NgR protein). In the homologous recombination vector, the altered portion of the NgR gene is flanked at its 5' and 3' ends by additional nucleic acid of the NgR gene to allow for homologous recombination to occur between the exogenous NgR gene carried by the vector and an endogenous NgR gene in an embryonic stem cell. The additional flanking NgR nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NgR gene has homologously recombined with the endogenous NgR gene are selected (see *e.g.*, Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. See *e.g.*, Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A Practical Approach, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr. Opin. Biotechnol.* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For

- 34 -

a description of the cre/loxP recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate

5 expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

10 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal

15 of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

20 Antisense

Also provided by the invention are antisense polynucleotides that recognize and hybridize to NgR polynucleotides. Full-length and fragment antisense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that

25 specifically recognize and hybridize to NgR RNA (as determined by sequence comparison of DNA encoding NgR to DNA encoding other known molecules). Identification of sequences unique to NgR encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the

30 desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be

- 35 -

performed. Antisense polynucleotides are particularly relevant to regulating expression of NgR by those cells expressing NgR mRNA.

Antisense oligonucleotides, or fragments of a nucleotide sequence set forth in SEQ ID NO:1, 3, 13 or sequences complementary or homologous thereto, derived
5 from the nucleotide sequences of the present invention encoding NgR are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. In specific aspects, antisense
10 nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire NgR coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a NgR protein of SEQ ID NO:2, 4 or 14 or antisense nucleic acids complementary to a NgR nucleic acid sequence of SEQ ID NOs:1, 3 or
15 13 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding NgR. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the protein coding region of human
20 NgR corresponds to the coding region SEQ ID NO:1, 3 or 13). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding NgR. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

25 Antisense oligonucleotides are preferably directed to regulatory regions of a nucleotide sequence of SEQ ID NO:1, 3, 13 or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like. Given the coding strand sequences encoding NgR disclosed herein (*e.g.*, SEQ ID NO:1, 3 or 13), antisense nucleic acids of the invention can be designed according
30 to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NgR mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding

- 36 -

or noncoding region of NgR mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NgR mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be
5 constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the
10 antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-
15 carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-
20 mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-
25 diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

30 The antisense nucleic acid molecules of the invention (preferably oligonucleotides of 10 to 20 nucleotides in length) are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA

- 37 -

and/or genomic DNA encoding a NgR protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. Suppression of NgR expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant NgR expression. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix.

Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine or cholesterol moieties at their 5' end.

An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.*, (1987) *Nucleic Acids Res.* 15, 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, (1987) *Nucleic Acids Res.* 15, 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, (1987) *FEBS Lett.* 215, 327-330).

The NgR sequences taught in the present invention facilitate the design of novel transcription factors for modulating NgR expression in native cells and animals,

- 38 -

and cells transformed or transfected with NgR polynucleotides. For example, the Cys₂-His₂ zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular NgR target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal *et al.*, (1999) *Proc. Natl. Acad. Sci. USA* 96, 2758-2763; Liu *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 5525-5530; Greisman *et al.* (1997) *Science* 275, 657-661; Choo *et al.*, (1997) *J. Mol. Biol.* 273, 525-532). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to ensure specificity for a particular sequence (Segal *et al.*, (1999), above). The artificial zinc finger repeats, designed based on the promoter of NgR sequences, are fused to activation or repression domains to promote or suppress NgR expression (Liu *et al.*, (1997), above). The promoter of NgR may be obtained by standard methods known to one of ordinary skill in the art with the disclosure contained herein and knowledge of the NgR sequence. Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 3616-3620. Such proteins and polynucleotides that encode them, have utility for modulating NgR expression *in vivo* in both native cells, animals and humans; and/or cells transfected with NgR-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 96, 9521-9526); Wu *et al.*, (1995) *Proc. Natl. Acad. Sci. USA* 92, 344-348). The present invention contemplates methods of

- 39 -

designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate NgR expression in cells (native or transformed) whose genetic complement includes these sequences.

5 **Ribozymes and PNA moieties**

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes,
10 described in Haselhoff and Gerlach (1988) *Nature* 334, 585-591) can be used to catalytically cleave NgR mRNA transcripts to thereby inhibit translation of NgR mRNA. A ribozyme having specificity for a NgR-encoding nucleic acid can be designed based upon the nucleotide sequence of a NgR DNA disclosed herein (*i.e.*, SEQ ID NOs:1, 3 or 13). For example, a derivative of a Tetrahymena L-19 IVS RNA
15 can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a NgR-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, NgR mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*,
20 (1993) *Science* 261, 1411-1418.

Alternatively, NgR gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NgR (*e.g.*, the NgR promoter and/or enhancers) to form triple helical structures that prevent transcription of the NgR gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6:
25 569-584; Helene. *et al.*, (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *BioEssays* 14, 807-815.

In various embodiments, the nucleic acids of NgR can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate
30 backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, (1996) *Bioorg. Med. Chem. Lett.* 4, 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in

which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard
5 solid phase peptide synthesis protocols as described in Hyrup *et al.*, (1996) above; Perry-O'Keefe *et al.*, (1996) *Proc. Natl. Acad. Sci. USA* 93,14670-14675.

PNAs of NgR can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or
10 inhibiting replication. PNAs of NgR can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.*, (1996), above; Perry-O'Keefe (1996), above).

15 In another embodiment, PNAs of NgR can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NgR can be generated that may combine the advantageous properties of PNA and DNA. Such
20 chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), above). The synthesis of PNA-DNA chimeras can be
25 performed as described in Hyrup (1996), above and Finn *et al.* (1996) *Nucleic Acids Res.* 24, 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl) amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucleic Acids Res.* 17,
30 973-988). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996), above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment

and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg. Med. Chem. Lett.* 5:1119-1124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating
5 transport across the cell membrane (see Letsinger *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86, 6553-6556; Lemaitre *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (see, *e.g.*, Krol *et al.*, (1988) *Biotechniques* 6,
10 958-976) or intercalating agents (see, *e.g.*, Zon (1988) *Pharm. Res.* 5, 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

Automated sequencing methods can be used to obtain or verify the nucleotide
15 sequence of NgR. The NgR nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid
20 molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule,
25 starting at the point of the mutation.

Polypeptides

The invention also provides purified and isolated mammalian NgR polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human NgR
30 polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:14. Another preferred embodiment is a mouse NgR polypeptide comprising the amino acid sequence of NgR3, as set forth in SEQ ID NO:4.

One aspect of the invention pertains to isolated NgR proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-NgR antibodies. Preferably, fragments of NgR proteins comprise at least one biological activity of NgR. In one embodiment, native NgR proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NgR proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a NgR protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, at least 50% or at least 45% identity and/or homology to the preferred polypeptide of the invention. In addition, the invention embraces polypeptides having the consensus sequence shown in SEQ ID NO:6, shown in Table 5) excluding the previously characterized NgR ("NgR1"), and polypeptides comprising at least about 90% of the consensus sequence.

The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment (Dayhoff, in ATLAS OF PROTEIN SEQUENCE AND

- 43 -

STRUCTURE, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference).

A determination of homology or identity is typically made by a computer homology program known in the art. An exemplary program is the Gap program
5 (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (*Adv. Appl. Math.*, 1981, 2, 482-489, which is incorporated herein by reference in its entirety). Employing the GAP software provided in the GCG program package, (see *Needleman and Wunsch* (1970) *J. Mol.*
10 *Biol.* 48, 443-453) the following settings for nucleic acid sequence comparison may be used: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NOs:1, 3 or 13.
15 BestFit was originally written for Version 1.0 by Paul Haeberli from a careful reading of the papers by Needleman and Wunsch (1970), above, and Smith and Waterman (1981), above. The following Bestfit settings for nucleic acid sequence comparison may be used: GAP creation penalty of 8.0 and GAP extension penalty of 2, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of
20 identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99%, with the CDS (encoding) part of the amino acid sequence shown in SEQ ID NOs:2, 4 or 14.

Alternatively, homology may be determined by hybridization analysis wherein a nucleic acid sequence is hybridized to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent
25 conditions. See *e.g.* Ausubel, *et al.*, (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention.

30 An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NgR protein is derived, or substantially free from

chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of NgR protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NgR protein having less than about 30% (by dry weight) of non-NgR protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-NgR protein, still more preferably less than about 10% of non-NgR protein, and most preferably less than about 5% non-NgR protein. When the NgR protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein having less than about 30% (by dry weight) of chemical precursors or non-NgR chemicals, more preferably less than about 20% chemical precursors or non-NgR chemicals, still more preferably less than about 10% chemical precursors or non-NgR chemicals, and most preferably less than about 5% chemical precursors or non-NgR chemicals.

Biologically active portions of a NgR protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the NgR protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, 4 or 14 that include fewer amino acids than the full length NgR proteins, and exhibit at least one activity of a NgR protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the NgR protein. A biologically active portion of a NgR protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a NgR protein of the present invention may contain at least one of the features that is conserved between the NgR proteins (*e.g.*, a conserved cysteine as the N-terminus of the mature protein, four conserved cysteines

- 45 -

in the N-terminus before a leucine-rich region, four conserved cysteines C-terminal with respect to a leucine repeat region, eight leucine-rich repeats, and a hydrophobic C-terminus). An alternative biologically active portion of a NgR protein may contain at least two of the above-identified domains. Another biologically active portion of a
5 NgR protein may contain at least three of the above-identified domains. Yet another biologically active portion of a NgR protein of the present invention may contain at least four of the above-identified domains.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one
10 or more of the functional activities of a native NgR protein.

In an embodiment, the NgR protein has an amino acid sequence shown in SEQ ID NO:2, 4 or 14. In other embodiments, the NgR protein is substantially homologous to SEQ ID NO:2, 4 or 14 and retains the functional activity of the protein of SEQ ID NO:2, 4 or 14, yet differs in amino acid sequence due to natural allelic variation or
15 mutagenesis, as described in detail below.

Accordingly, in another embodiment, the NgR protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or SEQ ID NO:14 and retains the functional activity of the NgR proteins of SEQ ID NO:2, 4 or 14.

20 Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of NgR polypeptides are embraced by the invention.

25 The invention also embraces variant (or analog) NgR polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a NgR amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the NgR amino acid sequence. Insertional variants with additional residues at either or both termini
30 can include, for example, fusion proteins and proteins including amino acid tags or labels.

Insertion variants include NgR polypeptides wherein one or more amino acid residues are added to a NgR acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature NgR products, *i.e.*, NgR products wherein leader or signal sequences are removed, with additional amino
5 terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. NgR products with an additional methionine residue at position -1 (Met⁻¹-NgR) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met⁻²-Lys⁻¹-NgR). Variants of NgR with
10 additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

Polypeptide Variants

The invention also embraces NgR variants having additional amino acid
15 residues which result from use of specific expression systems.

As used herein, a NgR "chimeric protein" or "fusion protein" comprises a NgR polypeptide operatively linked to a non-NgR polypeptide. A "NgR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to NgR, whereas a
20 "non-NgR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not homologous to the NgR protein, *e.g.*, a protein that is different from the NgR protein and that is derived from the same or a different organism. Within a NgR fusion protein the NgR polypeptide can correspond to all or a portion of a NgR protein. In one embodiment, a NgR fusion protein comprises at least one biologically active portion of a NgR protein. In another embodiment, a NgR
25 fusion protein comprises at least two biologically active portions of a NgR protein. In yet another embodiment, a NgR fusion protein comprises at least three biologically active portions of a NgR protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the NgR polypeptide and the non-NgR polypeptide are fused in-frame to each other. The non-NgR polypeptide can be fused to the N-
30 terminus or C-terminus of the NgR polypeptide.

For example, in one embodiment a NgR fusion protein comprises a NgR domain operably linked to the extracellular domain of a second protein. Such fusion

- 47 -

proteins can be further utilized in screening assays for compounds which modulate NgR activity (such assays are described in detail below).

For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the
5 desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide.

In another embodiment, the fusion protein is a NgR protein containing a heterologous signal sequence at its N-terminus. For example, the native NgR signal sequence (*i.e.*, amino acids 1-30 of SEQ ID NO:2 and amino acids 1-40 of SEQ ID
10 NO:4) can be removed and replaced with a signal sequence from another protein. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion NgR can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is a NgR-immunoglobulin fusion protein in which the NgR sequences comprising one or more domains are fused to
15 sequences derived from a member of the immunoglobulin protein family. The NgR-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between NgR ligand and a NgR protein on the surface of a cell, to thereby suppress NgR-mediated signal transduction *in vivo*. NgR-immunoglobulin fusion proteins can be
20 used to affect the bioavailability of a NgR cognate ligand. Inhibition of the NgR ligand/NgR interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the NgR-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NgR antibodies in a subject, to
25 purify NgR ligands, and in screening assays to identify molecules that inhibit the interaction of NgR with NgR ligand.

A NgR chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional
30 techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and

- 48 -

enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can
5 subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel *et al.* (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A NgR-encoding nucleic acid can be cloned into such an expression vector such that the fusion
10 moiety is linked in-frame to the NgR protein.

Variants resulting from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of NgR is/are fused to another polypeptide.

15 In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a NgR polypeptide are removed. Deletions can be effected at one or both termini of the NgR polypeptide, or with removal of one or more non-terminal amino acid residues of NgR. Deletion variants, therefore, include all fragments of a NgR polypeptide.

20 The invention also embraces polypeptide fragments of the sequence set forth in SEQ ID NO:2, 4 or 14 wherein the fragments maintain biological (*e.g.*, ligand binding and/or intracellular signaling) immunological properties of a NgR polypeptide. Fragments comprising at least 4, 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of SEQ ID NO:2, 4 or 14 are contemplated by the invention. Preferred
25 polypeptide fragments display antigenic properties unique to, or specific for, human NgR and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of NgR
30 polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a NgR polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the

invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 2, 3, or 4 below.

Table 1.

5	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁	G, R, M	
	X ₂	A, D, C	
	X ₃	V, T	
10	X ₄	N, P, S	
	X ₅	E, A, S	
	X ₆	nothing, K	nothing
	X ₇	V, M, P	
	X ₈	T, V	V
15	X ₉	Q, P	Q
	X ₁₀	Q, A	Q
	X ₁₁	Q, H, N	
	X ₁₂	G, N	N
	X ₁₃	L, F	F
20	X ₁₄	Q, A, S	
	X ₁₅	A, S	
	X ₁₆	V, I	
	X ₁₇	V, T, E, L	
	X ₁₈	S, G	
25	X ₁₉	L, I	
	X ₂₀	A, E, V, P	
	X ₂₁	A, S, D	
	X ₂₂	S, T	
	X ₂₃	Q, E	

- 50 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₂₄	IVL	
	X ₂₅	Q,H	Q
	X ₂₆	N,G	N
	X ₂₇	R,L	
5	X ₂₈	T,G,R,S	
	X ₂₉	F,L,T,H	
	X ₃₀	L,V	L
	X ₃₁	Q,R,P	
	X ₃₂	Q,P,A	P
10	X ₃₃	G,A	G
	X ₃₄	H,T,S	
	X ₃₅	S,G,R	
	X ₃₆	P,S,A	
	X ₃₇	C, nothing	nothing
15	X ₃₈	R, nothing	nothing
	X ₃₉	A,N	
	X ₄₀	M,L	
	X ₄₁	V,L,T	
	X ₄₂	T, I	T
20	X ₄₃	L,I	
	X ₄₄	Y,F,H	
	X ₄₅	N,V	N
	X ₄₆	I,L	
	X ₄₇	T,S,A	
25	X ₄₈	F,Y,T,R	
	X ₄₉	A,H,Y,D	

- 51 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₅₀	P,A	P
	X ₅₁	N,S,G,A	
	X ₅₂	T,A	T
	X ₅₃	E,R,T	
5	X ₅₄	G,H	
	X ₅₅	F,L	
	X ₅₆	V,Q,H	
	X ₅₇	H,A,L	
	X ₅₈	E,Q	E
10	X ₅₉	G,S	G
	X ₆₀	R,A	R
	X ₆₁	Q,H	Q
	X ₆₂	R, H	H
	X ₆₃	T,S	
15	X ₆₄	L,V	L
	X ₆₅	A,E,D	
	X ₆₆	E,D,A	
	X ₆₇	Q,H	Q
	X ₆₈	V,E,G	
20	X ₆₉	K,R	
	X ₇₀	H,Q	
	X ₇₁	A,S,T	
	X ₇₂	Y,H	
	X ₇₃	Y,D	Y
25	X ₇₄	K,R	
	X ₇₅	G,Q	

- 52 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₇₆	S,Q	S
	X ₇₇	A,S,E	
	X ₇₈	P,G	P
	X ₇₉	A,G,P	
5	X ₈₀	G,N	
	X ₈₁	I,V,L	
	X ₈₂	G,R	
	X ₈₃	H,V,A	
	X ₈₄	S,A	S
10	X ₈₅	D,E	
	X ₈₆	H,S,A	
	X ₈₇	I,L	
	X ₈₈	E,L,Q	
	X ₈₉	Y,H,A	
15	X ₉₀	Q,P	Q
	X ₉₁	D, N	
	X ₉₂	I, L, T	
	X ₉₃	V,A,R	
	X ₉₄	V,A,G	
20	X ₉₅	S,T	S
	X ₉₆	K,R	
	X ₉₇	L,I	L
	X ₉₈	W,R,S	
	X ₉₉	S,L	
25	X ₁₀₀	L,V	L
	X ₁₀₁	G,T,P	

- 53 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₀₂	Q,P,E	
	X ₁₀₃	G,H,R	
	X ₁₀₄	I,T,V,A	
	X ₁₀₅	V,G,H	
5	X ₁₀₆	N,S	
	X ₁₀₇	E,G,Q	
	X ₁₀₈	Q,R	
	X ₁₀₉	L,V	
	X ₁₁₀	Q,A	
10	X ₁₁₁	W,G,H	
	X ₁₁₂	H,R,P	
	X ₁₁₃	K,A,H	
	X ₁₁₄	H,R	
	X ₁₁₅	D,G	
15	X ₁₁₆	H,R,S,G	
	X ₁₁₇	T,M	
	X ₁₁₈	T,I	
	X ₁₁₉	F,Y	
	X ₁₂₀	N,A	
20	X ₁₂₁	S,N	
	X ₁₂₂	T,A,S	
	X ₁₂₃	E,S,A	
	X ₁₂₄	Q,P	
	X ₁₂₅	G,T	
25	X ₁₂₆	D,E	D
	X ₁₂₇	C,A	

- 54 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₂₈	P,D	
	X ₁₂₉	V,G,P,R	
	X ₁₃₀	A,S	
	X ₁₃₁	E,Q	Q
5	X ₁₃₂	F,Y	F
	X ₁₃₃	G,A,D	
	X ₁₃₄	A,P	
	X ₁₃₅	D,A,V	
	X ₁₃₆	G,D	
10	X ₁₃₇	A,E	
	X ₁₃₈	S,P	
	X ₁₃₉	E,A	
	X ₁₄₀	L,F	
	X ₁₄₁	R,Q	
15	X ₁₄₂	R,K	R
	X ₁₄₃	R,K	R
	X ₁₄₄	F,A	
	X ₁₄₅	G,V	
	X ₁₄₆	A,D,E	
20	X ₁₄₇	T,P	
	X ₁₄₈	A,V,S	
	X ₁₄₉	T,S,L	
	X ₁₅₀	E,G,P,Q	
	X ₁₅₁	L,E,R	
25	X ₁₅₂	R,L	R
	X ₁₅₃	G, D	

- 55 -

X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
X ₁₅₄	Q,H,A	
X ₁₅₅	Q,R	
X ₁₅₆	K,R	
X ₁₅₇	L,A,R	
5	X ₁₅₈	R
X ₁₅₉	V,A,E	
X ₁₆₀	E,A,N	
X ₁₆₁	F,L	F
X ₁₆₂	R,Q	
10	X ₁₆₃	N,A,G

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 2 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

20

Table 2**Conservative Substitutions I**

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Aliphatic Non-polar	G A P I L V

- 56 -

Polar - uncharged	C S T M N Q
Polar - charged	D E K R
Aromatic	H F W Y
Other	N Q D E

5 Alternatively, conservative amino acids can be grouped as described in
Lehninger, [BIOCHEMISTRY, Second Edition; Worth Publishers, Inc. NY, NY
(1975), pp.71-77] as set out in Table 3, immediately below.

Table 3

Conservative Substitutions II

10

	SIDE CHAIN CHARACTERISTIC	AMINO ACID
	Non-polar (hydrophobic)	
	A. Aliphatic:	A L I V P
	B. Aromatic:	F W
15	C. Sulfur-containing:	M
	D. Boderline:	G
	Uncharged-polar	
	A. Hydroxyl:	S T Y
	B. Amides:	N Q
	C. Sylfhydryl:	C
20	D. Boderline:	G
	Positively Charged (Basic):	K R H
	Negatively Charged (Acidic):	D E

As still another alternative, exemplary conservative substitutions are set out in
25 Table 4, below.

Table 4

Conservative Substitutions III

	Original Residue	Exemplary Substitution
	Ala (A)	Val, Leu, Ile
30	Arg (R)	Lys, Gln, Asn
	Asn (N)	Gln, His, Lys, Arg
	Asp (D)	Glu
	Cys (C)	Ser
	Gln (Q)	Asn
35	Glu (E)	Asp

- 57 -

	His (H)	Asn, Gln, Lys, Arg
	Ile (I)	Leu, Val, Met, Ala, Phe,
	Leu (L)	Ile, Val, Met, Ala, Phe
5	Lys (K)	Arg, Gln, Asn
	Met (M)	Leu, Phe, Ile
	Phe (F)	Leu, Val, Ile, Ala
	Pro (P)	Gly
	Ser (S)	Thr
	Thr (T)	Ser
10	Trp (W)	Tyr
	Tyr (Y)	Trp, Phe, Thr, Ser
	Val (V)	Ile, Leu, Met, Phe, Ala

In addition, amino acid residues that are conserved among family members of the NgR proteins of the present invention, as indicated by the alignment presented herein, are also predicted to be particularly unamenable to alteration. For example, NgR proteins of the present invention can contain at least one domain that is a typically conserved region in NgRs. Examples of these conserved domains include, *e.g.*, leucine-rich repeat domain. Amino acid residues that are not conserved or are only semi-conserved among members of the NgR proteins may be readily amenable to alteration.

Full-length NgRs have an LRR region characterized by the amino acid consensus sequence shown in SEQ ID NO: 19. At least some full-length NgRs also include a CT signaling (CTS) domain and a GPI domain.

The NgR domain designations used herein are defined as follows:

Domain	hNgR1 SEQ ID: 5	mNgR1 SEQ ID NO:17	hNgR2 SEQ ID: 2	hNgR3 SEQ ID: 14	mNgR3 SEQ ID: 4
Signal Seq.	1-26	1-26	1-30	—	1-40
LRRNT	27-56	27-56	31-59	—	41-69
LRR1	57-81	57-81	60-82	5-27	70-92
LRR2	82-105	82-105	83-106	28-51	93-106
LRR3	106-130	106-130	107-131	52-76	106-141
LRR4	131-154	131-154	132-155	77-100	142-165

- 58 -

5	LRR5	155-178	155-178	156-179	101-124	166-189
	LRR6	179-202	179-202	180-203	125-148	190-213
	LRR7	203-226	203-226	204-227	149-172	214-237
	LRR8	227-250	227-250	228-251	173-196	238-261
	LRRCT	260-309	260-309	261-310	206-255	271-320
	CTS (CT Signaling)	310-445	310-445	311-395	256-396	321-438
	GPI	446-473	456-473	396-420	370-392	439-462

10

In some embodiments of the invention, the above domains are modified. Modification can be in a manner that preserves domain functionality. Modification can include addition, deletion or substitution of certain amino acids. Exemplary modifications include conservative amino acid substitutions. Preferably such

15 substitutions number 20 or fewer per 100 residues. More preferably, such substitutions number 10 or fewer per 100 residues. Further exemplary modifications include addition of flanking sequences of up to five amino acids at the N terminus and/or C terminus of one or more of the domains.

20 In some embodiments, the isolated nucleic acid molecule encodes a polypeptide at least about 70%, 80%, 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2, 4 or 14.

Mutations can be introduced into SEQ ID NOS:1, 3 or 13 by standard techniques, e.g., site-directed mutagenesis and PCR-mediated mutagenesis. Conservative amino acid substitutions can be made at one or more amino acid residues

25 predicted to be non-essential. Alternatively, mutations can be introduced randomly along a NgR coding sequence. This can be accomplished, e.g., by saturation mutagenesis. The resulting mutants can be screened for NgR biological activity. Biological activities of NgR may include but are not limited to: (1) protein:protein interactions, e.g., with other NgRs or other cell-surface proteins involved in Nogo-related signaling; (2) complex formation with a NgR ligand; (3) binding to an anti-

30 NgR antibody.

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues or organs. Similarly, the invention further embraces NgR polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol or polypropylene glycol. Variants that display ligand binding properties of native NgR and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

Chemically modified NgR polypeptide compositions in which the NgR polypeptide is linked to a polymer are included within the scope of the present invention. The polymer may be water soluble to prevent precipitation of the protein in an aqueous environment, such as a physiological environment. Suitable water-soluble polymers may be selected from the group consisting of, for example, polyethylene glycol (PEG), monomethoxypolyethylene glycol, dextran, cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, polypropylene glycol homopolymers, a polypropylene oxide/ethylene oxide copolymer polyoxyethylated polyols (*e.g.* glycerol) and polyvinyl alcohol. The selected polymer is usually modified to have a single reactive group, such as an active ester for acylation or an aldehyde for alkylation, so that the degree of polymerization may be controlled. Polymers may be of any molecular weight, and may be branched or unbranched, and mixtures of such polymers may also be used. When the chemically modified NgR polymer is destined for therapeutic use, pharmaceutically acceptable polymers will be selected for use.

When the polymer is to be modified by an acylation reaction, the polymer should have a single reactive ester group. Alternatively, if the polymer is to be modified by reductive alkylation, the polymer should have a single reactive aldehyde

group. A preferred reactive aldehyde is polyethylene glycol propionaldehyde, which is water stable, or mono Cl-CIO alkoxy or aryloxy derivatives thereof (see U.S. Patent No. 5,252,714, incorporated by reference herein in its entirety).

Pegylation of NgR polypeptides may be carried out by any of the pegylation reactions known in the art, as described, for example, in the following references: *Focus on Growth Factors* 3, 4-10 (1992); EP 0 154 316; and EP 0 401 384 (each of which is incorporated by reference herein in its entirety). Preferably, the pegylation is carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer). A preferred water-soluble polymer for pegylation of polypeptides such as NgR is polyethylene glycol (PEG). As used herein, "polyethylene glycol" is meant to encompass any of the forms of PEG that have been used to derivatize other proteins, such as mono (Cl-CIO) alkoxy- or aryloxy-polyethylene glycol.

Chemical derivatization of NgR polypeptides may be performed under any suitable conditions used to react a biologically active substance with an activated polymer molecule. Methods for preparing pegylated NgR polypeptides will generally comprise the steps of (a) reacting the polypeptide with polyethylene glycol, such as a reactive ester or aldehyde derivative of PEG, under conditions whereby NgR polypeptide becomes attached to one or more PEG groups, and (b) obtaining the reaction products. It will be apparent to one of ordinary skill in the art to select the optimal reaction conditions or the acylation reactions based on known parameters and the desired result.

Pegylated and other polymer:NgR polypeptides may generally be used to treat conditions that may be alleviated or modulated by administration of the NgR polypeptides described herein. However, the chemically-derivatized polymer:NgR polypeptide molecules disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the nonderivatized molecules. The NgR polypeptides, fragments thereof, variants and derivatives, may be employed alone, together, or in combination with other pharmaceutical compositions. The cytokines, growth factors, antibiotics, antiinflammatories and/or chemotherapeutic agents as is appropriate for the indication being treated.

The present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient or medium.

5 Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil and cocoa butter.

10 Variants that display ligand binding properties of native NgR and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

15 With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences which encode NgR from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is
20 incorporated herein by reference in its entirety.

For example, DNA that encodes NgR may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the NgR gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in
25 accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.* (1989) above.

A nucleic acid molecule comprising any of the NgR nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the
30 nucleotide sequences provided herein. See U.S. Patent Nos. 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that

- 62 -

sequence has not been previously purified and is present only in a single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase-mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, METHODS IN ENZYMOLOGY 152 Academic Press, San Diego, CA, which is incorporated herein by reference in its entirety.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

15 **Antibodies**

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for NgR or fragments thereof. Preferred antibodies of the invention are human antibodies which are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')₂, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind NgR polypeptides exclusively (*i.e.*, are able to distinguish NgR polypeptides from other known NgR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between NgR and such polypeptides).

The antigenic peptide of NgR comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2, 4 or 14 and encompasses an epitope of

NgR such that an antibody raised against the peptide forms a specific immune complex with NgR. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. Preferred epitopes
5 encompassed by the antigenic peptide are regions of NgR that are located on the surface of the protein, *e.g.*, hydrophilic regions.

It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and,
10 in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* in ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments
15 of the NgR polypeptides of the invention are also contemplated, provided that the antibodies are specific for NgR polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by injection with the native
20 protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed NgR protein or a chemically synthesized NgR polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*,
25 aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum* or similar immunostimulatory agents. If desired, the antibody molecules directed against NgR can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such
30 as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition," as used herein, refers to a population of antibody molecules that contain only one species

- 64 -

of an antigen binding site capable of immunoreacting with a particular epitope of NgR. A monoclonal antibody composition thus typically displays a single binding affinity for a particular NgR protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular NgR protein, or derivatives, fragments,
5 analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler and Milstein (1975) *Nature* 256, 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor *et al.*, (1983) *Immunol. Today* 4, 72) and the EBV hybridoma
10 technique to produce human monoclonal antibodies (see Cole *et al.*, (1985) in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote *et al.*, (1983) *Proc. Natl. Acad. Sci. USA* 80, 2026-2030) or by transforming human B-cells with Epstein Barr
15 Virus *in vitro* (see Cole *et al.*, (1985), above).

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a NgR protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of Fab expression libraries (see *e.g.*, Huse *et al.*, (1989) *Science* 246, 1275-1281) to allow rapid and effective
20 identification of monoclonal F_{ab} fragments with the desired specificity for a NgR protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced
25 into the antibody framework to modulate affinity or immunogenicity. Antibody fragments that contain the idiotypes to a NgR protein may be produced by techniques known in the art including, but not limited to: (i) an F(ab')₂ fragment produced by pepsin digestion of an antibody molecule; (ii) an Fab fragment generated by reducing the disulfide bridges of an F(ab')₂ fragment; (iii) an Fab fragment generated by the
30 treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Additionally, recombinant anti-NgR antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be

5 produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.*,

10 (1988) *Science* 240, 1041-1043; Liu *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 3439-3443; Liu *et al.*, (1987) *J. Immunol.* 139, 3521-3526; Sun *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 214-218; Nishimura *et al.*, (1987) *Cancer Res.* 47, 999-1005; Wood *et al.*, (1985) *Nature* 314, 446-449; Shaw *et al.*, (1988) *J. Natl. Cancer Inst.* 80, 1553-1559; Morrison (1985) *Science* 229, 1202-1207; Oi *et al.*, (1986)

15 *BioTechniques* 4, 214; U.S. Patent. No. 5,225,539; Jones *et al.*, (1986) *Nature* 321, 552-525; Verhoeyan *et al.*, (1988) *Science* 239, 1534; and Beidler *et al.*, (1988) *J. Immunol.* 141, 4053-4060.

In a preferred embodiment of the invention a portion of a NgR is joined to an Fc portion of an antibody to form a NgR/Fc fusion protein. Preferably, the Ig fusion

20 protein is soluble. The NgR/Fc fusion protein may be formed by recombinant techniques as described above. In one embodiment, a portion of a NgR including the entire amino acid sequence of NgR except the C-terminal hydrophobic region is fused to an Fc portion of an antibody. In preferred embodiments, the NgR is a human NgR and the Fc is also human. More preferably, the human Fc portion is derived from an

25 IgG antibody. In other embodiments, the N-terminal signal sequence is omitted. Such antibodies are useful in binding Nogo to prevent Nogo signaling through the NgR.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a

30 specific embodiment, selection of antibodies that are specific to a particular domain of a NgR protein is facilitated by generation of hybridomas that bind to the fragment of a NgR protein possessing such a domain. Antibodies that are specific for one or more

domains within a NgR protein, *e.g.*, domains spanning the above-identified conserved regions of NgRs, or derivatives, fragments analogs or homologs thereof, are also provided herein.

Anti-NgR antibodies may be used in methods known within the art relating to the localization and/or quantitation of a NgR protein (*e.g.*, for use in measuring levels of the NgR protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for NgR proteins, or derivatives, fragments analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-NgR antibody (*e.g.*, monoclonal antibody) can be used to isolate NgR by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-NgR antibody can facilitate the purification of natural NgR from cells and of recombinantly produced NgR expressed in host cells. Moreover, an anti-NgR antibody can be used to detect NgR protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the NgR protein. Anti-NgR antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering

- 67 -

to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Another aspect of the invention is directed to anti-idiotypic antibodies and
5 anti-anti-idiotypic antibodies. An anti-idiotypic antibody is an antibody that recognizes determinants of another antibody (a target antibody). Generally, the anti-idiotypic antibody recognizes determinants of the antigen-binding site of the target antibody. Typically, the target antibody is a monoclonal antibody. An anti-idiotypic antibody is generally prepared by immunizing an animal (particularly, mice) of the same species
10 and genetic type as the source of the target monoclonal antibody, with the target monoclonal antibody. The immunized animal mounts an immune response to the idiotypic determinants of the target monoclonal antibody and produces antibodies against the idiotypic determinants of the target monoclonal antibody. Antibody-producing cells, such as splenic cells, of the immunized animal may be used
15 to generate anti-idiotypic monoclonal antibodies. Furthermore, an anti-idiotypic antibody may also be used to immunize animals to produce anti-anti-idiotypic antibodies. These immunized animals may be used to generate anti-anti-idiotypic monoclonal antibodies using standard techniques. The anti-anti-idiotypic antibodies may bind to the same epitope as the original, target monoclonal antibody used to
20 prepare the anti-idiotypic antibody. The anti-anti-idiotypic antibodies represent other monoclonal antibodies with the same antigen specificity as the original target monoclonal antibody.

If the binding of the anti-idiotypic antibody with the target antibody is inhibited by the relevant antigen of the target antibody, and if the anti-idiotypic antibody induces
25 an antibody response with the same specificity as the target antibody, it mimics the antigen of the target antibody. Such an anti-idiotypic antibody is an "internal image anti-idiotype" and is capable of inducing an antibody response as if it were the original antigen. (Bona and Kohler (1984) ANTI-IDIOTYPIC ANTIBODIES AND INTERNAL IMAGE, IN MONOCLONAL AND ANTI-IDIOTYPIC ANTIBODIES: PROBES FOR RECEPTOR STRUCTURE
30 AND FUNCTION, Venter J.C. *et al.* (Eds), Alan R. Liss, New York, NY, pp 141-149, 1984). Vaccines incorporating internal image anti-idiotype antibodies have been shown to induce protective responses against viruses, bacteria, and parasites (Kennedy

et al., (1986) 232, 220-223; 1047; McNamara *et al.*, (1985) *Science* 226, 1325-1326). Internal image anti-idiotypic antibodies have also been shown to induce immunity to tumor related antigens (Raychauhuri *et al.*, (1986) *J. Immunol.* 137, 1743-1749; Raychauhuri *et al.*, (1987) *J. Immunol.* 139, 3902-3910; Bhattacharya-Chatterjee *et al.*, (1987) *J. Immunol.* 139, 1354-1360; Bhattacharya-Chatterjee *et al.*, (1988) *J. Immunol.* 141, 1398-1403; Herlyn. *et al.* (1989) *Intern. Rev. Immunol.* 4, 347-357; Chen *et al.* (1990) *Cell Imm. Immunother. Cancer* 351-359; Herlyn *et al.*, (1991) *in vivo* 5, 615-624; Furuya *et al.* (1992) *AntiCancer Res.* 12, 27-32; Mittelman, A. *et al.* (1992) *Proc. Natl. Acad. Sci., USA* 89, 466-470; Durrant. *et al.*, (1994) *Cancer Res.* 54, 4837-4840; Mittelman. *et al.* (1994) *Cancer Res.* 54, 415-421; Schmitt. *et al.* (1994) *Hybridoma* 13, 389-396; Chakrobarty. *et al.* (1995) *J. Immunother.* 18, 95-103; Chakrobarty. *et al.* (1995) *Cancer Res.* 55, 1525-1530; Foon, K. A. *et al.* (1995) *Clin. Cancer Res.* 1, 1205-1294; Herlyn *et al.* (1995) *Hybridoma* 14, 159-166; Sclebusch *et al.* (1995) *Hybridoma* 14, 167-174; Herlyn. *et al.* (1996) *Cancer Immunol Immunother.* 43, 65-76).

Anti-idiotypic antibodies for NgR may be prepared, for example, by immunizing an animal, such as a mouse, with a immunogenic amount of a composition comprising NgR2 (SEQ ID NO:2), NgR3 (SEQ ID NOs:4 or 14), or immunogenic portion thereof, containing at least one antigenic epitope of NgR. The composition may also contain a suitable adjuvant, and any carrier necessary to provide immunogenicity. Monoclonal antibodies recognizing NgR may be prepared from the cells of the immunized animal as described above. A monoclonal antibody recognizing an epitope of NgR is then selected and used to prepare a composition comprising an immunogenic amount of the anti-NgR monoclonal antibody. Typically, a 25 to 200 μ g dose of purified anti-NgR monoclonal would be sufficient in a suitable adjuvant.

Animals may be immunized 2-6 times at 14 to 30 day intervals between doses. Typically, animals are immunized by any suitable route of administration, such as intraperitoneal, subcutaneous, intravenous or a combination of these. Anti-idiotypic antibody production may be monitored during the immunization period using standard immunoassay methods. Animals with suitable titers of antibodies reactive with the target monoclonal antibodies may be reimmunized with the monoclonal antibody used as the immunogen three days before harvesting the antibody producing cells.

- 69 -

Preferably, spleen cells are used, although other antibody producing cells may be selected. Antibody-producing cells are harvested and fused with myeloma cells to produce *Hybridomas*, as described above, and suitable anti-idiotypic antibody-producing cells are selected.

- 5 Anti-anti-idiotypic antibodies are produced by another round of immunization and *Hybridoma* production by using the anti-idiotypic monoclonal antibody as the immunogen.

- Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of NgR), diagnostic purposes to detect or quantitate NgR, and
10 purification of NgR. Therefore, kits comprising an antibody of the invention for any of the purposes described herein are also comprehended.

Kits

- The present invention is also directed to kits, including pharmaceutical kits.
15 The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well appropriate controls, such as positive and/or negative controls. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.
20 For example, the kit can comprise: a labeled compound or agent capable of detecting NgR protein or mRNA in a biological sample; means for determining the amount of NgR in the sample; and means for comparing the amount of NgR in the sample with a standard. The compound or agent can be packaged in a suitable container.

25 Screening Assays

- The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a NgR polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein NgR polypeptides are
30 immobilized and cell-based assays. Identification of binding partner compounds of NgR polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with NgR normal and aberrant biological activity.

- 70 -

The invention also provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules (*e.g.*, molecules of less than 1,000 Daltons) or other drugs) that bind to NgR proteins or have a stimulatory or inhibitory effect on, for example, NgR expression or NgR activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a NgR protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12, 145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90, 6909; Erb *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91, 11422; Zuckermann *et al.* (1994) *J. Med. Chem* 37, 2678; Cho *et al.*, (1993) *Science* 261, 1303; Carrell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2059; Carell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2061; and Gallop *et al.*, (1994) *J. Med. Chem* 37, 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *BioTechniques* 13, 412-421), or on beads (Lam (1991) *Nature* 354, 82-84), on chips (Fodor (1993) *Nature* 364, 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, above), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89, 1865-1869) or on phage (Scott and Smith (1990) *Science* 249, 386-390; Devlin (1990) *Science* 249, 404-406; Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 6378-6382; Felici (1991) *J. Mol. Biol.* 222, 301-310; Ladner, above).

30

1. Cell-based Assays

The invention also provides cell-based assays to identify binding partner compounds of a NgR polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a NgR polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the NgR polypeptide. In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the NgR protein or biologically active portion thereof.

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a NgR protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NgR protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NgR protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NgR protein or a biologically active portion thereof, on the cell surface with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the test compound to preferentially bind to NgR or a biologically active portion thereof as compared to the known compound.

- 72 -

Determining the ability of the test compound to modulate the activity of NgR or a biologically active portion thereof can be accomplished, for example, by determining the ability of the NgR protein to bind to or interact with a NgR target molecule. As used herein, a "target molecule" is a molecule with which a NgR protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a NgR protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A NgR target molecule can be a non-NgR molecule or a NgR protein or polypeptide of the present invention. In one embodiment, a NgR target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound NgR molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with NgR. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant NgR products, NgR variants, or preferably, cells expressing such products. Binding partners are useful for purifying NgR products and detection or quantification of NgR products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of NgR, especially those activities involved in signal transduction.

25

2. Cell-free Assays

(a) Direct binding:

The invention includes several assay systems for identifying NgR binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a NgR polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the NgR polypeptide. Identification of the compounds that bind the NgR polypeptide can be achieved by isolating the NgR

30

- 73 -

polypeptide/binding partner complex and separating the binding partner compound from the NgR polypeptide. An additional step of characterizing the physical, biological and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the NgR polypeptide/binding partner complex is isolated using an antibody immunospecific for either the NgR

5 polypeptide or the candidate binding partner compound.

In still other embodiments, either the NgR polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the NgR

10 polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG[®] tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

15

(b) Immobilized NgR

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized NgR polypeptide, or a biologically active fragment thereof with a candidate binding partner compound and (b)

20 detecting binding of the candidate compound to the NgR polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of NgR is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead or a chromatographic resin, as well as non-covalent, high affinity interactions such as

25 antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Binding of a test compound to NgR, or interaction of NgR with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge

30 tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, and not by way of limitation, GST-NgR fusion proteins or GST-target fusion proteins can be

- 74 -

adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or NgR protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, and the complexes determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of NgR binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either NgR or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated NgR or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NgR or target molecules, but which do not interfere with binding of the NgR protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NgR trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NgR or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NgR or target molecule.

Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, (v) determining the activity of the NgR, as well as other techniques well known and routinely practiced in the art.

- 75 -

Determining the activity of the target molecule, for example, may be accomplished by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a NgR-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

(c) Competition experiments

In yet another embodiment, the assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the test compound to preferentially bind to NgR or biologically active portion thereof as compared to the known compound.

In yet another embodiment, the cell-free assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the NgR protein to modulate the activity of a NgR target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of NgR. In the case of cell-free assays comprising the membrane-bound form of NgR, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of NgR is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-

- 76 -

propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

Modulators

5 Agents that modulate (*i.e.*, increase, decrease, or block) NgR activity or expression may be identified by incubating a putative modulator with a cell containing a NgR polypeptide or polynucleotide and determining the effect of the putative modulator on NgR activity or expression. The selectivity of a compound that modulates the activity of NgR can be evaluated by comparing its effects on NgR to its
10 effect on other NgR compounds. Selective modulators may include, for example, antibodies and other proteins, peptides or organic molecules which specifically bind to a NgR polypeptide or a NgR-encoding nucleic acid. Modulators of NgR activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant NgR activity is involved. NgR polynucleotides, polypeptides and
15 modulators may be used in the treatment of such diseases and conditions associated with demyelination. NgR polynucleotides and polypeptides, as well as NgR modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations
20 including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the NgR polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A
25 modulator that increases binding between the NgR polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the NgR polypeptide and the binding partner compound is described as an inhibitor.

In another embodiment, modulators of NgR expression may be identified in a
30 method wherein a cell is contacted with a candidate compound and the expression of NgR mRNA or protein in the cell is determined. The level of expression of NgR mRNA or protein in the presence of the candidate compound is compared to the level

- 77 -

of expression of NgR mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NgR expression based on this comparison. For example, when expression of NgR mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NgR mRNA or protein expression. Alternatively, when expression of NgR mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NgR mRNA or protein expression. The level of NgR mRNA or protein expression in the cells can be determined by methods described herein for detecting NgR mRNA or protein.

High Throughput Screening

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, etc.) of a NgR polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate NgR receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the NgR polypeptide.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either NgR or nucleic acid molecules encoding NgR, comprising contacting NgR, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds NgR or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, Ausubel *et al.* (Eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in

- 78 -

its entirety. The NgR proteins, for example, can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, (1993) *Cell* 72, 223-232; Madura *et al.*, (1993) *J. Biol. Chem.* 268, 12046-12054; Bartel *et al.*, (1993) *BioTechniques* 14, 920-924; Iwabuchi *et al.*, (1993) *Oncogene* 8, 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with NgR ("NgR-binding proteins" or "NgR-bp") and modulate NgR activity. Such NgR-binding proteins are also likely to be involved in the propagation of signals by the NgR proteins as, for example, upstream or downstream elements of the NgR pathway.

Other assays may be used to identify specific ligands of a NgR receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, (1989) *Nature* 340, 245-246, and Fields *et al.*, (1994) *Trends Genet.* 10, 286-292, both of which are incorporated herein by reference. The two-hybrid system is a genetic assay based on the modular nature of most transcription factors used for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the

- 79 -

noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a NgR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal. The compounds to be screened include (which may include compounds that are suspected to bind NgR, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biological or chemical origin.

10 The function of the NgR gene product is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay is useful to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a NgR receptor, or fragment thereof, a fusion polynucleotide encoding both a NgR receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used.

15 In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even

20 necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein.

25 One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a

30 test ligand which binds the target protein, than in the absence of a ligand. Binding of

- 80 -

the ligand to the target protein can be determined by any method which distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.* (1997) *Anal. Chem.* 69:1683-1691, incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I , ^{35}S , ^{32}P , ^{33}P , ^3H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The NgR polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between NgR and its substrate caused by the compound being tested.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of NgR comprising contacting NgR with a compound, and determining whether the compound modifies activity of NgR. The activity in the presence of the test compared is measured to the

- 81 -

activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample
5 lacking the test compound, the compound will have inhibited activity.

The present invention is particularly useful for screening compounds by using NgR in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate NgR activity), but are not limited to, extracellular, intracellular, biologic or chemical origin.
10 The NgR polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Nogo-R and its substrate caused by the
15 compound being tested.

The activity of NgR polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands. Alternatively, the activity of the NgR can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides,
20 neurotransmitters, nucleotides, lipids, odorants and photons. Alternatively, the activity of the NgR can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of NgR activity may alter a NgR receptor function, such as a binding property of a receptor or an activity. In various embodiments of the method, the assay
25 may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [³⁵S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca²⁺ concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (*e.g.*, using [³H]-arachidonic acid) and an assay for
30 extracellular acidification rates, as well as other binding or function-based assays of NgR activity that are generally known in the art. NgR activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those

- 82 -

skilled in the art. Biological activities of NgR receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of NgRs known in the art. Non-limiting examples of NgR activities include transmembrane signaling of various forms, which
5 may involve phosphatidylinositol (PI) association and/or the exertion of an influence over PI; another exemplary activity of NgRs is the binding of accessory proteins or polypeptides that differ from known GPI proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural NgR receptor ligands, peptide and non-peptide allosteric effectors of NgR receptors, and peptides that may
10 function as activators or inhibitors (competitive, uncompetitive and non-competitive) (e.g., antibody products) of NgR receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule
15 libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, ENZYME ASSAYS: A PRACTICAL APPROACH, Eisenthal and
20 Danson (Eds.), 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use
25 of radiolabelled ligands in HTS binding assays for drug discovery (see Williams (1991) *Med. Res. Rev.*, 11, 147-184; Sweetnam *et al.*, (1993) *J. Nat. Prod.* 56, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see
30 Hodgson (1992) *Bio/Technology* 10, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg *et al.* (1992) *Trends Pharmacol. Sci.* 13, 95-98), yeast (Pausch (1997) *Trends Biotechnol.* 15, 487-494), several kinds of insect cells (Vanden
5 Broeck (1996) *Int. Rev. Cytol.* 164, 189-268), amphibian cells (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt *et al.* (1997) *Eur. J. Pharmacol.* 334, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

10 In preferred embodiments of the invention, methods of screening for compounds which modulate NgR activity comprise contacting test compounds with NgR and assaying for the presence of a complex between the compound and NgR. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed
15 label is a measure of the ability of the particular compound to bind to NgR.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to NgR is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with NgR and washed. Bound NgR is then
20 detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed NgR can be used for HTS binding assays in
25 conjunction with its defined ligand. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I , ^3H , ^{35}S or ^{32}P , by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.* (1994) *Drug Dev. Res.* 33, 373-398; Rogers (1997) *Drug Discov. Today* 2, 156-160). Radioactive
30 ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate

- 84 -

- bound ligand from unbound ligand (Williams (1991) *Med. Res. Rev.* 11, 147-184; Sweetnam *et al.* (1993) *J. Nat. Prod.* 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama (1998) *Curr. Opin. Drug Disc. Dev.* 1, 85-91 Bossé *et al.* (1998) *J. Biomol. Screening* 3, 285-292). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers (1997) *Drug Discov. Today* 2, 156-160; Hill (1998) *Curr. Opin. Drug Disc. Dev.* 1, 92-97).
- 10 Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch (1997) *Trends in Biotechnol.* 15, 487-494); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy *et al.* (1998) *Cur. Opin. Drug Disc. Dev.* 1, 192-199). Fluorescence changes can also be used to
- 15 monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder *et al.* (1996) *J. Biomol. Screening* 1, 75-80). Melanophores prepared from *Xenopus laevis* show a ligand-dependent change in pigment organization in response to heterologous NgR activation; this response is adaptable to HTS formats
- 20 (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.
- 25 Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to transduce the signal for the binding of Nogo in membranes prepared from these cells through the putative GPI anchor. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the
- 30 radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca^{2+} concentration or membrane potential in permanently transfected CHO cells containing

- 85 -

each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (*e.g.*, Pausch (1997), above).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to NgR receptors. In one example, the NgR receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the NgR receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the NgR receptor and its binding partner. Another contemplated assay involves a variation of the di-hybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms that are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Cane *et al.*, *Science* (1998) 282, 63-68. Combinatorial libraries are composed of

- 86 -

large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide,
5 protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers (1997) *Curr. Opin. Biotechnol.* 8, 701-707. Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate
10 activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other
15 than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified NgR gene.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the
20 invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with NgR. Radiolabeled competitive binding studies are described in Lin *et al.*, (1997) *Antimicrob. Agents Chemother.* 41, 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

25 In other embodiments of the invention, the polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are
30 incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of

- 87 -

polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors.

- 5 In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing.
- 10 This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

- Alternatively, compounds may be identified which exhibit similar properties to the ligand for the NgR of the invention, but which are smaller and exhibit a longer half
- 15 time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid
- 20 randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

- Comparison of the protein sequence of the present invention with the
- 25 sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of NgR can be
- 30 designed.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Compositions and Pharmaceutical Compositions

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral or parenteral administration. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's PHARMACEUTICAL SCIENCES, 16th ed., (1980) Osol, A (Ed.), which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include oral and parenteral (*e.g.*, intravenous, intradermal, subcutaneous, inhalation, transdermal (topical), transmucosal and rectal administration). Solutions or suspensions used for parenteral, intradermal or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic

- 89 -

acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a NgR protein or anti-NgR antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of

- 90 -

preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They
5 can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.
10 Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel or corn starch; a lubricant
15 such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable
20 propellant, *e.g.*, a gas such as carbon dioxide or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile
25 salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with
30 conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

- 91 -

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811. It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent No. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Patent No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91, 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack or dispenser together with instructions for administration.

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to
5 500 mg/kg of body weight of the compound can be administered. Therapy is typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

Another aspect of the present invention is the use of the NgR nucleotide sequences disclosed herein for identifying homologs of the Nogo-R, in other animals,
10 including but not limited to humans and other mammals and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more
15 preferably at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with NgR sequences can be identified.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described
20 herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson *Science* (1992) 256, 808-813, which is incorporated herein by reference in its entirety.

25 The present invention also encompasses a method of agonizing (stimulating) or antagonizing a NgR natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a
30 mammal with an agonist or antagonist of the protein of the present invention comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize NgR-associated functions.

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996, and International patent publication number WO 96/22976, published August 1, 1996, both of which are
5 incorporated herein by reference in their entirety, including any drawings, figures or tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition
10 of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

15 Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection
20 methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell
25 composition. For example, toxicity studies can be carried out in a suitable animal model as follows: (1) the compound is administered to mice (an untreated control mouse should also be used); (2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and (3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes
30 versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

- 94 -

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, (1993) *J. Am. Vet. Med. Assoc.* 202:229-249). Representative
5 animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the
10 less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness.
15 Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

NgR mRNA transcripts have been found in the brain and heart. SEQ ID NOs: 1 and/or, 3 will, as detailed above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize NgR and
20 for compounds with potential utility in treating disorders including CNS disorders (*e.g.*, stroke) and degenerative disorders such as those associated with demyelination.

For example, NgR receptor activation may mediate the prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, *e.g.*, Donovan *et al.*, (1997) *J. Neurosci.* 17, 5316-5326; Turgeon *et al.*, (1998) *J. Neurosci.* 18, 6882-6891; Smith-Swintosky *et al.*, (1997) *J. Neurochem.* 69,
25 1890-1896; Gill *et al.*, (1998) *Brain Res.* 797, 321-327; Suidan *et al.*, (1996) *Semin. Thromb. Hemost.* 22, 125-133.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on NgR activity (*e.g.*, NgR gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, a disease condition such as a demyelination disorder) associated with aberrant NgR activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NgR protein, expression of NgR nucleic acid or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum (1996) *Clin. Exp. Pharmacol. Physiol.* 23, 983-985 and Linder (1997) *Clin. Chem.* 43, 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2

- 96 -

(NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of NgR protein, expression of NgR nucleic acid, or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a NgR modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NgR (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase NgR gene expression, protein levels or upregulate NgR activity, can be monitored in clinical trials of subjects exhibiting decreased NgR gene

expression, protein levels, or downregulated NgR activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NgR gene expression, protein levels, or downregulate NgR activity, can be monitored in clinical trials of subjects exhibiting increased NgR gene expression, protein levels, or
5 upregulated NgR activity. In such clinical trials, the expression or activity of NgR and, preferably, other genes that have been implicated in, for example, a disease or disorder, can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, genes, including NgR, that are modulated in cells by treatment
10 with an agent (*e.g.*, compound, drug or small molecule) that modulates NgR activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on demyelination disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of NgR and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a
15 gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced by one of the methods as described herein or by measuring the levels of activity of NgR or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state
20 may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate
25 identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a NgR protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the NgR
30 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the NgR protein, mRNA or genomic DNA in the pre-administration sample with the NgR protein, mRNA or genomic DNA in the post

administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NgR to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration
5 of the agent may be desirable to decrease expression or activity of NgR to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods
10 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
15 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to; (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous
20 insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous recombination (see, *e.g.*, Capecchi (1989) *Science* 244, 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction
25 between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner.
30 Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art
5 include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a
10 disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic
15 agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

20 Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a
25 protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity.
30 Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the

- 100 -

agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described
5 herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

10 Gene Therapy

Mutations in the NgR gene that result in loss of normal function of the NgR gene product underlie NgR human disease states. The invention comprehends gene therapy to restore NgR activity to treat those disease states. Delivery of a functional NgR gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors,
15 and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson (1998) *Nature*, supplement to 392(6679):25-20. For additional reviews of gene therapy technology see Friedmann (1989) *Science* 244, 1275-1281; Verma (1990) *Sci. Am.* 68-84; and Miller (1992)
20 *Nature* 357, 455-460. Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the activity of, NgR will be useful in treating disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of NgR.

The present invention provides for both prophylactic and therapeutic methods
25 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
30 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR

- 101 -

peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous recombination (see, *e.g.*, Capecchi (1989), above); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or

NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity. Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figure. Such modifications are intended to fall within the scope of the appended claims.

The following Table 5 contains the sequences of exemplary polynucleotides and polypeptides of the invention.

TABLE 5

The following DNA sequence NgR2 <SEQ ID NO. 1> was identified in humans:

```

5  ATGCTGCCCCGGGCTCAGGCGCCTGCTGCAAGCTCCCGCCTCGGCCTGCCTCCTGCTGATG
   CTCTGGCCCTGCCCCTGGCGGGCCCCAGCTGCCCCATGCTCTGCACCTGCTACTCATCC
   CCGCCACCGTGAGCTGCCAGGCCAACAACTTCTCCTCTGTGCCGCTGTCCCTGCCACCC
   AGCACTCAGCGACTCTTCCTGCAGAACAACTCATCCGCACGCTGCGGCCAGGCACCTTT
10  GGGTCCAACCTGCTCACCTGTGGCTCTTCTCCAACAACCTCTCCACCATCTACCCGGGC
   ACTTTCCGCCACTTGCAAGCCCTGGAGGAGCTGGACCTCGGTGACAACCGGCACCTGCGC
   TCGCTGGAGCCCGACACCTTCCAGGGCCTGGAGCGGCTGCAGTCGCTGCATTTGTACCGC
   TGCCAGCTCAGCAGCCTGCCCGGCAACATCTCCGAGGCCTGGTCAGCCTGCAGTACCTC
   TACCTCCAGGAGAACAGCCTGCTCCACCTACAGGATGACTTGTTGCGGGACCTGGCCAAC
15  CTGAGCCACCTCTTCTCCACGGGAACCGCCTGCGGCTGCTCACAGAGCACGTGTTTCGC
   GGCCTGGGCAGCCTGGACCGGCTGCTGCTGCACGGGAACCGGCTGCAGGGCGTGACCGC
   GCGGCCTTCCGCGGCCTCAGCCGCCTACCATCCTCTACCTGTTCAACAACAGCCTGGCC
   TCGCTGCCCCGGCGAGGCGCTCGCCGACCTGCCCTCGCTCGAGTTCTGCGGCTCAACGCT
   AACCCTGGGCGTGCGACTGCCGCGCGCGGCCGCTCTGGGCCTGGTTCCAGCGCGCGCGC
20  GTGTCCAGTCCGACGTGACCTGCGCCACCCCCCGAGCGCCAGGGCCGAGACCTGCGC
   GCGCTCCGCGAGGCCGACTTCCAGGCGTGTCGCCCGCGCGCACCCACGCGGCCGGGCGAGC
   CGCGCCCGCGGCAACAGCTCCTCAACACCTGTACGGGGTGGCCGAGGCCGGGGCGCCC
   CCAGCCGATCCCTCCACCTCTACCGAGATCTGCCTGCCGAAGACTCGCGGGGGCGCCAG
   GGCGGGGACGCGCCTACTGAGGACGACTACTGGGGGGGCTACGGGGGTGAGGACCAGCGA
25  GGGGAGCAGATGTGCCCCGGCGCTGCCTGCCAGGCGCCCCCGGACTCCCGAGGCCCTGCG
   CTCTCGGCCGGGCTCCCCAGCCCTCTGCTTTGCCCTCCTGCTCCTGGTGCCCCACCACCTC

```

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1:

```

30  M L P G L R R L L Q A P A S A C L L L M L L A L P L A A P S C
   P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S T
   Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N N
   L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E P
   D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L V
35  S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F L
   H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q G
   V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L A
   D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R A
   R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A C
40  P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P P
   A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W G
   G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S A
   G L P S P L L C L L L L V P H H L

```

- 104 -

The following DNA sequence Ngr3 <SEQ ID NO. 3> was identified in mouse:

5 ATGTCTTGGCAGTCTGGAACACAGTGACACAATCTCCCGTGCAGGCTGCTCAGGTCTCA
GGGTGCTGTGTGGAATTGCTGCTGTTGCTGCTCGCTGGAGAGCTACCTCTGGGTGGTGGT
TGTCCTCGAGACTGTGTGTGCTACCCTGCGCCCATGACTGTCAGCTGCCAGGCACACAAC
10 TTTGCTGCCATCCCGGAGGGCATCCCAGAGGACAGTGAGCGCATCTTCCTGCAGAACAAAT
CGCATCACCTTCTCCAGCAGGGCCACTTCAGCCCCGCCATCGTACCCCTCTGGATCTAC
TCCAACAACATCACTTTTCATTGCTCCCAACACCTTCGAGGGCTTTGTGCATCTGGAGGAG
CTAGACCTTGGAGACAACCGACAGCTGCGAACGCTGGCACCCGAGACCTTCCAAGGCCTG
GTGAAGCTTCACGCCCTCTACCTCTATAAGTGTGGACTGAGCGCCCTGCCCGCAGGCATC
15 TTTGGTGGCCTGCACAGCCTGCAGTATCTCTACTTGCAGGACAACCATATCGAGTACCTC
CAAGATGACATCTTTGTGGACCTGGTCAATCTCAGTCACTTGTTCCTCATGGTAACAAG
CTATGGAGCCTGGGCCAAGGCATCTTCGGGGCCTGGTGAACCTGGACCGGTTGCTGCTG
CATGAGAACCAGCTACAGTGGGTTACCCACAAGGCTTTCCATGACCTCCACAGGCTAACC
ACCTCTTTCTCTTCAACAACAGCCTCACTGAGCTGCAGGGTGACTGTCTGGCCCCCCTG
20 GTGGCCTTGGAGTTCCTTCGCCTCAATGGGAATGCTTGGGACTGTGGCTGCCGGGCACGT
TCCCTGTGGGAATGGCTGCGAAGGTTCCGTGGCTCTAGCTCTGCTGTCCCCTGCGCGACC
CCCGAGCTGCGGCAAGGCCAGGATCTGAAGCTGCTGAGGGTGGAGGACTTCCGGAACCTGC
ACAGGACCAGTGTCTCCTCACCAGATCAAGTCTCACACGCTTACCACCTCTGACAGGGCT
GCCCGCAAGGAGCACCATCCGTCCCATGGGGCCTCCAGGGACAAAGGCCACCCACATGGC
25 CATCCGCCTGGCTCCAGGTCAAGTTACAAGAAGGCAGGCAAGAACTGCACCAGCCACAGG
AACCGGAACCAGATCTCTAAGGTGAGCTCTGGGAAAGAGCTTACCGAACTGCAGGACTAT
GCCCCGACTATCAGCACAAAGTTAGCTTTGACATCATGCCACCGCACGACCCAAGAGG
AAGGGCAAGTGTGCTCGCAGGACCCCCATCCGTGCCCCAGTGGGGTGCAGCAGGCATCC
TCAGGCACGGCCCTTGGGGCCCCACTCCTGGCCTGGATACTGGGGCTGGCAGTCACTCTC
CGC

The following protein sequence <SEQ ID NO. 4> is deduced protein of SEQ ID NO:3:

30 MSWQSGTTVTQSPVQAAQVSGCCVELLLLLL
AGELPLGGGCPRDCVCYPAPMTVSCQAHNFA
AIPEGIPEDSERIFLQNNRITFLQQGHFSPA
MVTLWIYSNNITFIAPNTFEGFVHLEELDLG
DNRQLRTLAPETFQGLVKLHALYLYKCGLSA
LPAGIFGGGLHSLQYLYLQDNHIEYLQDDIFV
35 DLVNLSHLFLHGNKLWSLGGQIFRGLVNLDR
LLLHENQLQWVHHKAFHDLHRLTTLFLFNNS
LTELQGDCLAPLVALEFLRLNGNAWDCGCRA
RSLWEWLRRFRGSSSAVPCATPELRQGQDLK
LLRVEDFRNCTGPVSPHQIKSHTLTTSDDRAA
40 RKEHHPSHGASRDKGHPHGHPPGSRSGYKKA
GKNCTSHRNRNQISKVSSGKELTELQDYAPD
YQHKFSFDIMPTARPKRKGCARRTPIRAPS
GVQQASSGTALGAPLLAWILGLAVTLR

The following protein sequence <SEQ ID NO. 5> is NgR1 from humans:

MKRASAGGSRLLA WVLWLQAWQVAAPCPGA
 C
 5 CYNPKVTTSCPQQGLQAVPVGIPAAASQRI
 FLHG NRISHVPAASFRA CRNL TILWLHS NVL
 ARIDAA AFTGLALLEQLD LSDNAQLRSV DPA
 TFHGLGRLHTLHLDR CGLQELGPGLFRGLAA
 10 LQYLYLQDNALQALPDDTFRDLGNLTHLFLH
 GNRIS SVPERAFRGLHSLDRLL LHQNRVAHV
 HPHAFRDLGRLMTLYLFANNLSALPTEALAP
 LRALQYLR LNDNPWVCD CRARPLWAWLQKFR
 GSSEVP CSLPQRLAGRDLKRLAANDLQGCA
 VATGPYHPIWTGRATDEEPLGLPKCCQPDAA
 15 DKASVLEPGRPASAGNALKGRVPPGDSPPGN
 GSGPRHINDSPFGTLP GSAEPPLTAVRPEG S
 EPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
 GSGGGGTGDSEGS GALPSLTCSLTPLGLALV
 LWTVLGPC

20 The following amino acid sequence <SEQ ID NO:6> is a Consensus Sequence of
 NgR based on homology with NgR1

CPXXCXCYXXPXXXTXSCXXXXXXXXXXPX
 XXPXXXXXXXXFLXXNXIXXXXXXXXXXXFXXXXXXXXXLWX
 XSNXXXXXIXXXXXFXXXXXLEXLDLXDNXXLR
 25 XXXPXTFXGLXXLXLXLXXCXLXXLXXXXFX
 GLXXLQYLYLQXNXXXXXLXDDXFXDLXNLXH
 LFLHG NXXXXXXXXXXFRGLXXLDRLLLHXNX
 XXXVHX XAFXXLXRLXXLXLFXNXLXXLXXX
 XLAXLXXLXXLRLNXNXWXCXCRARXLWXX
 XXXRXSSSVXCXXPXXXXXGDXLXXLXXDX
 30 XXCXXXXXPXXPXXXXXXXXXXXXXXXXXXXXX
 XXXXXXXXXXXXXXXXXXXXXXXGXXXXXXXXXXXXX
 PXXXXSXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXRXXX
 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXLXXXXX
 35 XXXXXL

The following protein sequence <SEQ ID NO:7> is the 66 amino acid active domain
 of Nogo:

40 RIYKGV IQAIQKSDEGHFPFRAYLESEVAISE
 ELVQKYSNSALGHVNCTIKELRRLFLVDDL
 DSLK

The following protein sequence <SEQ ID NO:8> is the amino acid sequence of the mature NgR2:

5 CPMLCTCYSSPPTVSCQANNFSSVPLSLPPS
TQRLFLQNNLIRTLRPGTFGSNLLTLWLFSN
NLSTIYPGTFRHLQALEELDLGDNRH LRSLE
PDTFQGLERLQSLHLYRCQLSSLPGNIFRGL
VSLQYLYLQENSLHLQDDLFADLANLSHLF
LHGNRLRLLTEHVFRGLGSLDRLLLHGNRLQ
10 GVHRAAFRGLSRLTILYLFNNSLASLPGEAL
ADLPSLEFLRLNANPWACDCRARPLWAWFQR
ARVSSSDVT CATPPERQGRDLRALREADFQA
CPPAAPTRPGSRARGNSSSNHLYGVAEAGAP
PADPSTLYRDLPAEDSRGRQG GDAPTEDDYW
15 GGYGGEDQRGEQMCPGAACQAPPDSRGPALS
AGLPSPLLCLLLLVP HHL

The following protein sequence <SEQ ID NO:9> is the amino acid sequence of the mature NgR3:

20 CPRDCVCYPAPMTVSCQAHNF AAIPEGIPED
SERIFLQNNRITFLQQGHFSPAMVTLWIYSN
NITFIAPNTFEGFVHLEELDLGDNRQLRTLA
PETFQGLVKLHALYLYKCGLSALPAGIFGGL
HSLQYLYLQDNHIEYLQDDIFVDLVNLSHLF
LHGNKLWSLGGQGIFRGLVNLDRLLLHENQLQ
25 WVHHKAFHDLHRLTTLFLFNNSLTELQGDCL
APLVALEFLRLNGNAWDCGCRARSLWEWLR
FRGSSSAVPCATPELRQGQDLKLLRVEDFRN
CTGPVSPHQIKSHTLTTS DRAARKEHHP SHG
30 ASRDKGHPHGHPPGSRSGYK KAGKNCTSHRN
RNQISKVSSGKELTELQDYAPDYQHKFSFDI
MPTARPKRKKGKCARRTPIRAPSGVQQASSGT
ALGAPLLAWILGLAVTLR

35 The following amino acid sequence <SEQ ID NO:10> is a conserved cysteine motif
(Cysteine domain 1) of the NgR and homologs based on the Consensus Sequence:

CPXXCXC YXXPXXTXSC

The following amino acid sequence <SEQ ID NO:11> is a conserved cysteine motif (Cysteine domain 2) of the NgR and homologs based on the Consensus Sequence:

N X W X C X C R A R X L W X W X X X R X S S S X V X C X X P
X X X X G X D L X X L X X X D X X X C

5

The following amino acid sequence <SEQ ID NO:12> is a conserved Leucine-rich domain of the NgR and homologs based on the Consensus Sequence:

R X F L X X N X I X X X X X X F X X X X X X X L W X X S N
X X X X I X X X F X X X X X L E X L D L X D N X X L R X X X
P X T F X G L X X L X L X L X X C X L X X L X X X F X G L X
X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H L F L
H G N X X X X X X X X F R G L X X L D R L L H X N X X X X
V H X X A F X X L X R L X X L X L F X N X L X X L X X X X L A
X L X X L X X L R L

10

15

Unless otherwise indicated, X is any amino acid. For example, X where indicated may be no amino acid. Additional features of the invention will be apparent from the following Examples. Examples 1-5 are actual, while the remaining Examples are prophetic.

20

As shown by the following Examples, a gene encoding novel NgRs have been identified by computational analysis of DNA sequence data. The proteins encoded by NgR2 and NgR3 have a putative signal sequence, eight leucine-rich repeat domains in a conserved leucine-rich region (SEQ ID NO:12), a conserved cysteine-rich region (SEQ ID NO:10) N-terminal to the leucine-rich region, a second cysteine-rich domain (SEQ ID NO:11) C-terminal to the leucine-rich region, and a putative glycoposphatidylinositol-linkage (GPI-linkage) site. NgR2 and NgR3 differ from the previously identified NgR sequence. The NgR homologs, when compared to known NgRs, show a consensus sequence (SEQ ID NOs:6). The putative mature NgR2 and NgR3 are shown in Table 5 as SEQ ID NOs: 8 and 9, respectively.

25

30

Example 1: Tblastn query of the HTG database

The protein sequence for the human NgR (NgR1) (SEQ ID NO:5) was used to query the high throughput genomic (HTG) database the use of which is familiar to those skilled in the art. The HTG database is a part of GenBank, a comprehensive

35

- 108 -

NIH genetic sequence database, which includes an annotated collection of all publicly available DNA sequences (*Nucleic Acids Res.* (2000) 28, 15-8). The HTG database includes sequences obtained from genomic DNA. Within genomic DNA, genes are typically encoded by multiple segments of DNA called exons. Thus when one aligns a cDNA sequence (or a protein sequence encoded by a cDNA sequence) to a genomic sequence, the sequence will be broken up into segments depending on the number of exons in the gene.

The BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, (1990) *J. Mol. Biol.* 215, 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). The basic BLAST algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90, 5873-5877, which is incorporated herein by reference) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which

- 109 -

provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a NgR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a NgR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

To query the HTG database with the NgR protein sequence, we used a variation of the BLAST algorithm known as the tblastn program, which compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames (*J. Mol. Biol.* (1990) 215, 403-410; *Nucleic Acids Res.* (1997) 25, 3389-3402). The results of the tblastn search indicated the presence of genes in the database with a significant identity to the NgR. In addition to finding hits to genomic clones which contain the human and mouse NgR genes, we found hits to clones where the identity was not as high, but still very significant. Three human clones were found (Accession numbers: AC068514, AC016869, AC013606) with an e-value of 4e-43 and one mouse clone was found (Accession No. AC021768) with an e-value of 1e-78. The three human clones all appeared to encode the same gene, so further analysis was confined to AC013606.

Example 2: Prediction of the human NgR2 protein sequence (AC013606)

The human NgR protein sequence aligned with two regions of translated sequence from nucleotide sequence AC013606, indicating that the new gene was encoded by at least two exons. In order to define the complete gene, we used the computer program GENSCAN[™] (*J. Mol. Biol.* (1997) 268, 78-94) which can identify complete exon/intron structures of genes in genomic DNA. The gene prediction by GENSCAN[™] contained seven exons. By comparing these predicted exons to the NgR, it was concluded that the new human gene contains two of these exons and a part of another (containing the initiating methionine). The predicted cDNA (mRNA) encoded by these three exons was assembled from AC013606 (HTG11; deposited March 2000; length = 143899; GenBank release 118.0; SEQ ID NO: 15) by combining nucleotides from the three exons whose coordinates are: 123292-123322 (exon 1); 130035-130516 (exon 2); and 138589-139335 (exon 3). The sequence for this cDNA

- 110 -

sequence is SEQ ID NO:1 (nucleotide sequence of human NgR2; AC013606). The translation of this cDNA provides the protein sequence of human NgR2 (SEQ ID NO:2).

We used the protein sequence of human NgR2 as a query sequence against the human EST database. A number of hits of high significance were found indicating that the NgR2 mRNA is expressed in a number of tissues including fetal brain. Furthermore, two of these ESTs provided support for the exon structure that we deduced. One EST (Accession No: GB_EST19:AI346757) contains 565 nucleotides corresponding to amino acids 84-271 of the human NgR2 (SEQ ID No:4). This spans the second intron located between amino acids 171 and 172, and provides positive evidence for the splicing of exons 2 and 3 at the mRNA level. Another EST (GB_EST26:AI929019) contains 545 nucleotides, part of which corresponds to amino acids 1-75 of the human NgR2 (SEQ ID NO:2). This spans the first intron located between amino acids 10 and 11, and provides positive evidence for the splicing of exons 1 and 2 at the mRNA level.

Example 3: Prediction of the mouse NgR3 protein sequence (AC021768)

The human NgR protein sequence aligned with only one region of translated sequence from nucleotide sequence AC021768, indicating that most of the new mouse gene was encoded by one large exon. However, upon inspection, the protein encoded by this exon was missing an initiating methionine. In order to define the complete gene, we used the computer program GENSCAN as described above. The gene prediction by GENSCAN contained two exons; the large one found by visual inspection and a short one at the 5' end which provided an initiating methionine. The predicted cDNA (mRNA) encoded by these two exons was assembled from AC021768 (HTG14; deposited March 2000; length = 215980; GenBank release 118.0; SEQ ID NO: 16) by combining nucleotides from the two exons whose coordinates are: the complement of 164265-164325 (exon 1); and the complement of 155671-156992 (exon 2). The sequence for this cDNA sequence is SEQ ID NO:3 (nucleotide sequence of mouse NgR3; AC021768). The translation of this cDNA provides the protein sequence of mouse NgR3 (SEQ ID NO:4).

- 111 -

We used the protein sequence of mouse NgR3 as a query sequence against the mouse EST database. One hit of high significance was found indicating that the NgR2 mRNA is expressed in the heart. This EST (GB_EST20:AI428334) contains 463 nucleotides, part of which correspond to amino acids 45-193 of mouse NgR3 (SEQ ID NO:4).

Example 4: Similarity between the NgRs

An alignment between NgR1 and the two new receptors is shown in Fig. 1A-1B. The similarities between these proteins include:

- 10 (1) The SignalP program, which locates the signal sequence cleavage position, predicts a cleavage before the first conserved cysteine in all the proteins. Thus the mature protein in all cases will have a cysteine at the N-terminus.
- (2) All proteins contain eight Leucine Rich Repeats (LRR). LRRs are short sequence motifs present in a number of proteins with diverse functions and cellular
15 locations. These repeats are usually involved in protein-protein interactions. Each LRR is composed of a beta-alpha unit.
- (3) All three proteins contain a leucine rich repeat N-terminal domain (LRRNT), in which four cysteines are conserved. LRRs are often flanked by cysteine rich domains at both their N and C termini.
- 20 (4) All three proteins contain a LRR C-terminal domain (LRRCT). The LRRCTs of the three NgR proteins can be distinguished from those of other LRR containing proteins, by the pattern of typtophans and cysteines which are completely conserved in this domain.
- (5) All three proteins contain a conserved cysteine in the fourth LRR
25 domain.
- (6) All three proteins contain a conserved potential glycosylation site in the eighth LRR domain.
- (7) NgR2 and NgR3 have a hydrophobic C-terminus, as does NgR1, an indication that they probably also undergo a modification similar to NgR1, where a
30 GPI moiety is covalently linked to a C-terminal amino acid. This allows the protein to remain tethered to the cell.

Example 5: Preparation of Nogo Proteins

A Nogo binding assay was developed which utilizes a method widely used in examining semaphorin and ephrin axonal guidance function (Flanagan & Vanderhaeghen (1998) *Annu. Rev. Neurosci.* 21,3 09-345; Takahashi *et al.* (1999) *Cell* 99, 59-69). It involves fusing a secreted placental alkaline phosphatase (AP) moiety to the ligand in question to provide a biologically active receptor binding agent which can be detected with an extremely sensitive colorimetric assay. For Nogo, an expression vector is created encoding a signal peptide, a His6 tag for purification, AP, and the 66 amino acid active domain of Nogo. The fusion protein can be purified from the conditioned medium of transfected cells in milligram amounts. This protein is biologically active as a growth cone collapsing agent with an EC₅₀ of 1 nM.

Alternatively, a glutathione-S-transferase Nogo (GST-Nogo) fusion protein may be prepared. For GST-Nogo, an expression vector (*e.g.*, a pGEX vector) is created encoding a signal peptide, GST, and the 66 amino acid active domain of Nogo. GST-Nogo may be purified from the culture medium and used as a GST fusion protein, or GST may be cleaved from the Nogo portion of the fusion protein with an enzyme that recognizes the specific amino acid cleavage site engineered between the GST portion and the Nogo portion of the fusion protein. Such sites are part of the commercially available GST vectors. The specific cleavage sites and enzymes may be used in accordance with the Manufacturer's specifications.

It has been found that AP-Nogo is actually slightly more potent than GST-Nogo, perhaps because the protein is synthesized in a eukaryotic rather than a prokaryotic cell.

Binding of Nogo to immobilized NgR homologs may be performed in an ELISA-type assay in which AP-Nogo is allowed to react with an immobilized receptor homolog. Specificity of binding may be demonstrated in a competitive binding assay using increasing amounts of GST-Nogo in the type of assay to show a decreasing amount of binding of AP-Nogo (as judged in the colorimetric assay).

Example 6: Transfected COS Cell binding Assays

The homologs of the present invention may be used in transfection studies in COS cells to demonstrate binding of Nogo. Specifically, nucleotide sequences

- 113 -

encoding NgR2 and NgR3 may be transfected into COS cells using a suitable vector. Non-transfected COS-7 cells do not bind AP-Nogo. However, transfection of COS cells with nucleic acid sequences encoding NgRs will make them capable of binding Nogo. AP alone does not bind with any stable affinity to these transfected cells,
5 indicating that any affinity of Nogo for NgR2 or NgR3 would be due to the 66 amino acids derived from Nogo. Furthermore, specific affinity of Nogo for the NgR2 or NgR3 proteins may be tested in displacement of AP-Nogo assays using GST-Nogo. NgR2 and/or NgR3 may also bind homologs of Nogo, which may also be tested using this assay.

10

Example 7: Expression of NgR in Human Cell Lines using Northern Blot and a Random-Primed Probe

A Northern blot is purchased from a commercial source, or RNA samples from cells of interest are run on an agarose gel and blotted to a membrane using any of the
15 well known techniques for Northern blotting. The blot is probed with a fragment of NgR2 (SEQ ID NO:1) or NgR3 (SEQ ID NO:3). The probe is prepared from 50 ng of cDNA labeled by a random-primed method (Feinberg and Vogelstein (1983) *Anal. Biochem.* 132, 6-13). Hybridization is carried out at 68°C for 1 hour in ExpressHyb™ solution (Clontech, Cat. No. 8015-1) followed by washing with 2X
20 SSC/0.05% SDS at room temperature and two washes with 0.1X SSC/0.1% SDS at 50°C. Expression of NgR2 and/or NgR3 can be assessed by the presence of an appropriately sized band on the blot.

25

Example 8: Cloning of cDNA corresponding to NgRs

To obtain the full-length clone corresponding to NgR2 from a cDNA library, the following method may be used. A cDNA library is generated using standard methods from a tissue known to contain NgR2. Such a tissue was identified in Example 2. 1×10^6 plaque forming units from the cDNA library may be screened in
30 duplicate on OPTITRAN™ filters. The filters are hybridized with ^{32}P -labeled oligonucleotides that are generated from the ESTs corresponding to portions of NgR2. The hybridization reaction may consist of 400 mls plaque screen buffer (50mM Tris pH

- 114 -

7.5, 1M NaCl, 0.1% Sodium pyrophosphate, 0.2% Polyvinylpyrrolidone and 0.2% Ficoll) containing 10% Dextran sulfate and 100µg/ml tRNA and 80 pmol each ³²P-labeled oligonucleotide at 65°C overnight. The filters are washed twice with 2X SSC/1%SDS and twice with 1X SSC/1%SDS and exposed to film. Duplicate positives
5 are purified. DNA from each of these clones is analyzed by restriction enzyme digest followed by agarose gel electrophoresis and Southern blotting. The filters are hybridized to the ³²P-labeled oligonucleotides used for the original hybridization to confirm that inserts hybridize to the probe. The insert is then sequenced to confirm that it represents the cDNA for NgR2. Similar methods may be used to generate a
10 full-length clone corresponding to NgR3.

Alternatively, a full-length clone of NgR2 or NgR3 can be obtained by a person of ordinary skill in the art employing conventional PCR techniques.

Example 9: Hybridization Analysis to demonstrate NgR expression in the brain

15 The expression of NgR in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and sagittal rat brain cryosections (20 µm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are
20 processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at
25 room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A NgR-specific probe may be generated using PCR. Following PCR amplification, the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for
30 the sense strand of NgR, a cloned NgR fragment cloned in pBluescript II may be linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (*i.e.*, cRNA riboprobes) using the vector-borne T7 promoter and

- 115 -

commercially available T7 RNA polymerase. A probe specific for the antisense strand of NgR may also be readily prepared using the NgR clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3 promoter and cognate polymerase. The riboprobes may be labeled with [³⁵S]-UTP to yield a specific activity of about 0.40 x 10⁶ cpm/pmol for antisense riboprobes and about 0.65 x 10⁶ cpm/pmol for sense-strand riboprobes. Each riboprobe may be subsequently denatured and added (2 pmol/ml) to hybridization buffer which contains 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol. Microscope slides containing sequential brain cryosections may be independently exposed to 45 µl of hybridization solution per slide and silanized cover slips may be placed over the sections being exposed to hybridization solution. Sections are incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are then exposed to sense or antisense NgR-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment involving the exposure of slides to 20 µg/ml RNase A in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three high-stringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMax™ MR-1 film. After 13 days of exposure, the film is developed, and any significant hybridization signal is detected. Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with Kodak NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violet-stained cell bodies. Autoradiographic grains found between cell bodies indicate non-specific binding of the probe.

- 116 -

In some cases, such as using a probe to detect a NgR homolog in a heterologous species, in order to achieve optimal hybridization, it may be necessary to decrease the stringency conditions. Such conditions are well known to those of ordinary skill in the art and examples are provided above.

5 Expression of NgR in the brain provides an indication that modulators of NgR activity have utility for treating neurological disorders. Some other diseases for which modulators of NgR may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of NgR modulators, including NgR ligands and anti-NgR antibodies, to treat individuals having
10 such disease states is intended as an aspect of the invention.

Example 10: Northern Blot Analysis of NgR-RNA with a PCR-generated Probe

Northern blot hybridizations may be performed to examine the expression of NgR mRNA. A clone containing at least a portion of the sequence of SEQ ID NO:1
15 may be used as a probe. Vector-specific primers are used in PCR to generate a hybridization probe fragment for ³²P-labeling. The PCR is performed as follows:

	Mix:	1μl	NgR-containing plasmid
		2μl	fwd primer (10-50 pM)
20		2μl	rev primer (10-50 pM)
		10μl	10xPCR buffer (such as that provided with the enzyme, Amersham Pharmacia Biotech)
		1μl	10mM dNTP (such as #1 969 064 from Boehringer Mannheim)
		0.5μl	Taq polymerase (such as #27-0799-62, Amersham Pharmacia Biotech)
25		83.5μl	water

- 117 -

PCR is performed in a Thermocycler using the following program:

	94°C	5min		
5	94°C	1min		30 cycles
	55°C	1min		
	72°C	1min		
10	72°C	10min		

- The PCR product may be purified using QIAquick PCR Purification Kit (#28104) from Qiagen, and radiactively labeled with ^{32}P -dCTP (#AA0005/250, Amersham Pharmacia Biotech)) may be done by random priming using "Ready-to-go DNA Labeling Beads" (#27-9240-01) from Amersham Pharmacia Biotech.
- Hybridization is carried out on Human Multiple Tissue Northern Blot from Clontech as described in manufacturer's protocol, or on a Northern Blot prepared by running RNA samples from cells of interest on an agarose gel and blotting to a membrane using any of the known Northern blotting protocols. After exposure overnight on Molecular Dynamics Phosphor Imager screen (#MD146-814) bands of an appropriate size are visualized.

Example 11: Recombinant Expression of NgR in Eukaryotic Host Cells

A. Expression of NgR in Mammalian Cells

- To produce NgR protein, a NgR-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, a NgR-encoding sequence described in Table 4 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells using the transfection reagent FuGENE6™ (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing NgR are selected by growth in the presence of 100 µg/ml zeocin (Stratagene, LaJolla, CA). As an alternative to FuGENE6™, the expression vector may carry the gene for dihydrofolate reductase (dhfr) and selection of clones with methotrexate (MTX) drug pressure

- 118 -

allows for stable transformation of CHO cells. Optionally, NgR may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the NgR amino acid sequence, and the antisera is used to affinity purify Nogo-R.

- 5 The NgR also may be expressed in-frame with a tag sequence (*e.g.*, polyhistidine, hemagglutinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for NgR polypeptides, such as assays described below, do not require purification of NgR from the host cell.

10 **B. Expression of NgR in CHO cells**

- For expression of NgR in Chinese hamster ovary (CHO) cells, a plasmid bearing the relevant NgR coding sequence is prepared, using a vector which also bears the selectable marker dihydrofolate reductase (DHFR). The plasmid is transfected into CHO cells. Selection under MTX drug pressure allows for preparation of stable
- 15 transformants of a NgR (NgR2 or NgR3) in an expression plasmid carrying a selectable marker such as DHFR.

C. Expression of NgR in 293 cells

- For expression of NgR in mammalian cells 293 (transformed human, primary
- 20 embryonic kidney cells), a plasmid bearing the relevant NgR coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for
- 25 selection of stable transfectants. The forward primer for amplification of this NgR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the HindIII cloning site and nucleotides matching the NgR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning
- 30 and nucleotides corresponding to the reverse complement of the NgR sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII-XhoI* sites of the vector.

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAP™ transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-NgR peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by Western blots probed with anti-His, anti-Myc or anti-NgR peptide antibodies.

D. Transient Expression of Nogo-R in COS cells

For expression of the NgR in COS7 cells, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, for example, can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an *Xba*I restriction site for cloning, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1. The reverse primer is also determined by routine procedures and preferably contains 5'-extension of nucleotides which introduces a *Sal*I cloning site followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1.

The PCR consists of an initial denaturation step of 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the *Xba*I and *Sal*I sites of vector p3-CI. This construct is transformed into *E. coli* cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

NgR expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for

- 120 -

example, chromatography. Purified NgR is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C. NgR3 may also be expressed using this method and the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:13.

5

E. Expression of NgR in Insect Cells

For expression of NgR in a baculovirus system, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, 3 or 13 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively). The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively).

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (PharMingen, San Diego, CA). The pAcHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pAcHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of NgR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (1987) A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555.

- 121 -

In a preferred embodiment, pAcHLT-A containing NgR gene is introduced into baculovirus using the "BaculoGold™" transfection kit (Pharmingen, San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with ³⁵S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a NgR polypeptide in a Sf9 cells, a polynucleotide molecule having the nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13) can be amplified by PCR using the primers and methods described above for baculovirus expression. The NgR cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the NgR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

F. Expression of soluble forms of NgR2 and NgR3 as NgR-Ig fusion proteins.

To generate a NgR2-Ig fusion protein, standard methods may be used as described in the literature (*e.g.* Sanicola *et al.* (1997) *Proc. Natl. Acad. Sci. USA*. 94, 6238-6243). For example, a DNA fragment encoding NgR2 without the sequence encoding the hydrophobic C-terminus (GPI anchor signal) may be ligated to a DNA fragment encoding the Fc domain of IgG1 (which may be human IgG1), and the chimeric fragment may be cloned into an expression vector to generate a plasmid. The plasmid may then be transfected into Chinese hamster ovary cells to generate a stable cell line producing the fusion protein. The fusion protein is then purified from conditioned media using standard methods. For example, clarified conditioned media from the cell line may be loaded by gravity directly onto Protein A Sepharose. The column may then be washed with five column volumes each of PBS, PBS containing

- 122 -

0.5 M NaCl, and 25 mM sodium phosphate, 100 mM NaCl (pH 5.0). The bound protein may then be eluted with 25 mM NaH₂PO₄, 100 mM NaCl (pH 2.8) and immediately neutralized with 1/10 fraction volume of 0.5 M Na₂HPO₄ (pH 8.6).

Similar methods may be used to generate a NgR3-Ig fusion protein.

5

Example 12: Interaction Trap/Two-Hybrid System

In order to assay for NgR-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.* (1989) *Nature* 340, 245, which is incorporated herein by reference in its entirety. A protocol is published in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 1999, John Wiley & Sons, NY and Ausubel, F. M. *et al.* 1992, SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

15 A fusion of the nucleotide sequences encoding all or partial NgR and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (*i.e.*, pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7) from cDNA of potential NgR-binding proteins (for protocols on forming cDNA libraries, see Sambrook *et al.* 1989, MOLECULAR CLONING: A LABORATORY
20 MANUAL, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/NgR fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure
25 expression in host cells and lack of transcriptional activity. Yeast cells are transformed (ca. 10⁵ transformants/mg DNA) with both the NgR and library fusion plasmids according to standard procedure (Ausubel, *et al.*, 1992, SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is
30 incorporated herein by reference in its entirety). *In vivo* binding of DNA-BD/NgR with AD/library proteins results in transcription of specific yeast plasmid reporter genes (*i.e.*, *lacZ*, *HIS3*, *ADE2*, *LEU2*). Yeast cells are plated on nutrient-deficient

- 123 -

media to screen for expression of reporter genes. Colonies are dually assayed for β -galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) supplemented media (filter assay for β -galactosidase activity is described in Breeden *et al.*, (1985) *Cold Spring Harb. Symp. Quant. Biol.*, 50, 643, which is incorporated herein by reference in its entirety).

5 Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific NgR/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine

10 the identity of the NgR-binding protein.

Example 13: Antibodies to Nogo-R

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the NgR receptor, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be

15 found, for example, in Sambrook *et al.* (1989), above, and Harlow *et al.* (Eds.), ANTIBODIES A LABORATORY MANUAL; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988). In one embodiment, recombinant NgR polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate

20 the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of NgR (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of Nogo-R, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a

25 hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant NgR or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger

30 mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Limpet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with

- 124 -

Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of NgR antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to confirm the presence of antibodies that immunoreact with NgR. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize NgR. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, UT) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 µM sodium hypoxanthine, 0.4 µM aminopterin, 16 µM thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5×10^6 thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning, NY).

On days 2, 4, and 6 after the fusion, 100 µl of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to NgR. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-NgR antibodies are obtained.

- 125 -

B. Humanization of anti-NgR monoclonal antibodies

The expression pattern of NgR as reported herein and the potential of NgRs as targets for therapeutic intervention suggest therapeutic indications for NgR inhibitors (antagonists). NgR-neutralizing antibodies comprise one class of therapeutics useful as
5 NgR antagonists. Following are protocols to improve the utility of anti-NgR monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-NgR antibodies).

The principles of humanization have been described in the literature and are
10 facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest
15 with the constant domains of human antibody molecules. (See, *e.g.*, Morrison *et al.*, (1989) *Adv. Immunol.*, 44, 65-92). The variable domains of NgR-neutralizing anti-NgR antibodies are cloned from the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and
20 the resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned
25 into human antibody sequences. (See, *e.g.*, Jones *et al.*, (1986) *Nature* 321, 522-525; Riechmann *et al.*, (1988) *Nature* 332, 323-327; Verhoeyen *et al.*, (1988) *Science* 239, 1534-1536 ; and Tempest *et al.*, (1991) *Bio/Technology* 9, 266-271). If necessary, the β -sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding
30 domain of the original monoclonal antibody. (See Kettleborough *et al.*, (1991) *Protein Engin.* 4, 773-783; and Foote *et al.*, (1992) *J. Mol. Biol.* 224, 487-499).

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan (1991) *Mol. Immunol.* 28, 489-498 .

5 The foregoing approaches are employed using NgR-neutralizing anti-NgR monoclonal antibodies and the hybridomas that produce them to generate humanized NgR-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein NgR expression or ligand-mediated NgR signaling is detrimental.

10 **C. Human NgR-Neutralizing Antibodies from Phage Display**

Human NgR-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.* (1997) *Human Antibodies* 8, 155-168; Hoogenboom (1997) *TIBTECH* 15, 62-70; and Rader *et al.* (1997), *Curr. Opin. Biotechnol.* 8, 503-508, all of which are incorporated by reference. For example,
15 antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is
20 expressed in bacteria, and the library is screened for NgR-specific phage-antibodies using labeled or immobilized NgR as antigen-probe.

D. Human NgR-neutralizing antibodies from transgenic mice

Human NgR-neutralizing antibodies are generated in transgenic mice essentially
25 as described in Bruggemann *et al.* (1996) *Immunol. Today* 17, 391-397 and Bruggemann *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 455-458. Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a NgR composition using conventional immunization protocols. hybridomas are generated using B cells from the
30 immunized mice using conventional protocols and screened to identify hybridomas secreting anti-NgR human antibodies (*e.g.*, as described above).

Example 14: Assays to Identify Modulators of NgR Activity

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of NgR activity. Among the modulators that can be identified by these assays are natural ligand compounds of the receptor; synthetic
5 analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind NgR are useful for identifying NgR in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes,
10 and the like). Agonist and antagonist modulators are useful for up-regulating and down-regulating NgR activity, respectively, to treat disease states characterized by abnormal levels of NgR activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or *visa versa*).

15

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in NgR-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See,
20 *e.g.*, Sutherland *et al.*, (1968) *Circulation* 37, 279; Frandsen *et al.*, (1976) *Life Sciences* 18, 529-541; Dooley *et al.*, (1997) *J. Pharmacol. Exp. Therap.* 283, 735-41; and George *et al.*, (1997) *J. Biomol. Screening* 2, 235-40). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

25 Briefly, the NgR coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well
30 microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with

- 128 -

wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (*i.e.*, candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [¹²⁵I]-labeled cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) and fixed amounts of [¹²⁵I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of NgR modulating activity. Modulators that act as agonists of receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP levels. Agonists of receptors which couple to the G_{i/o} subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

B. Aequorin Assays

In another assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a NgR expression construct and a construct that encodes the photoprotein apoaequorin. In the presence of the cofactor coelenterazine, apoaequorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. (See generally, Cobbold, *et al.* "Aequorin measurements of cytoplasmic free calcium," *In*: McCormack J.G. and Cobbold P.H., eds., CELLULAR CALCIUM: A PRACTICAL APPROACH. Oxford:IRL Press (1991); Stables *et al.*, (1997) *Anal. Biochem.* 252, 115-26; and Haugland, HANDBOOK OF FLUORESCENT PROBES AND RESEARCH CHEMICALS. Sixth edition. Molecular Probes, Eugene, OR (1996)).

In one exemplary assay, NgR is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaequorin (Molecular Probes, Eugene, OR) into CHO cells

- 129 -

using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to serum-free MEM containing 5 µM coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate NgR modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 µl/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 µl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC₅₀ values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for assaying for modulators of NgR activity. Cells (e.g., CHO cells or COS 7 cells) are transiently co-transfected with both a NgR expression construct (e.g., NgR in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George *et al.* (1997) *J. Biomol. Screening* 2, 235-240; and Stratowa *et al.* (1995) *Curr. Opin. Biotechnol.* 6, 574-581). Luciferase

- 130 -

activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in
5 MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with both a NgR expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (Legally, CA). Transfections are
10 performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are
15 incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450
20 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3-20-fold stimulation of
25 luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of
30 receptor activity, and such assays can be employed to screen for modulators of NgR activity. For example, CHO cells stably transfected with a NgR expression vector are plated at a density of 4×10^4 cells/well in Packard black-walled, 96-well plates

- 131 -

5 specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3™
5 AM, Fluo-4™ AM, Calcium Green™ -1 AM, or Oregon Green™ 488 BAPTA-1 AM), each at a concentration of 4 μM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of
10 the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μM; positive control), or ATP (4 μM; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.*
15 (1998) *Drug Dev. Res.* 44,14-20). The F-stop for the detector camera is set at 2.5 and the length of exposure is 0.4 milliseconds. Basal fluorescence of cells is measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level is subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium
20 ionophore A23187 and ATP increase the calcium signal 200% above baseline levels. In general, activated NgRs increase the calcium signal at least about 10-15% above baseline signal.

E. [³⁵S]GTPγS Binding Assay

25 It is also possible to evaluate whether NgR signals through a G protein-mediated pathway. Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [³⁵S]-GTPγS in the presence and absence of candidate modulators provides another assay
30 for modulator activity. (See, *e.g.*, Kowal *et al.*, (1998) *Neuropharmacology* 37, 179-187.).

- 132 -

In one exemplary assay, cells stably transfected with a NgR expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE
5 buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (1 ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to
10 isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl_2 , 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C
15 are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM MgCl_2 , 1 mM EDTA, 120 mM NaCl, 10 μM GDP, and 0.2 mM ascorbate, at a concentration of 10-50 $\mu\text{g}/\text{ml}$. In a final volume of 90 μl , homogenates are incubated with varying concentrations of candidate modulator compounds or 100 μM GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 μl guanosine
20 5'-O-(3 ^{35}S)thio triphosphate (NEN, 1200 Ci/mmol; ^{35}S]-GTP γS), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl_2 , at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with
25 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM MgCl_2 . Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of ^{35}S]-GTP γS is measured in the presence of 100 μM GTP and subtracted from the total. Compounds are selected that modulate the amount of ^{35}S]-GTP γS binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in
30 ^{35}S]-GTP γS binding. This response is blocked by antagonists.

F. ^3H Arachidonic Acid Release

- 133 -

The activation of NgRs may also potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of NgR activity. (See, *e.g.*, Kanterman *et al.*, (1991) *Mol. Pharmacol.* 39,364-369.) For example, CHO cells that are stably transfected with a NgR expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [³H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 µCi/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10 µM ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [³H]-arachidonic acid. This potentiation is blocked by antagonists.

G. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of NgR activity are assayed by monitoring extracellular changes in pH induced by the test compounds (see, *e.g.*, Dunlop *et al.* (1998) *J. Pharmacol. Toxicol. Meth.* 40, 47-55). In one embodiment, CHO cells transfected with a NgR expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4×10^5 cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 µg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO₂ for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 µg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 µl/minute.

- 134 -

Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

Example 15: mNgR3 does not bind hNogo-A(1055-1120)

To functionally test the mouse NgR3 (hereinafter, mNgR3) for its ability to bind hNogo-A(1055-1120), a cDNA expression vector for a myc epitope-tagged mNgR3 protein was created. The mouse NgR3 cDNA was amplified by PCR from mouse adult brain cDNA, from the signal sequence to the stop codon, and ligated into the pSecTag2 vector such that the vector encodes a signal sequence followed by a myc tag followed by the mature mNgR3 sequence. This plasmid was transfected into COS7 cells, and expression of a myc-tagged protein of the predicted size was verified by immunoblot analysis. Alkaline phosphatase-hNogo-A(1055-1120) binding studies and myc immunohistology were conducted as described (Fournier et al., supra).

The cells expressing mNgR3 express the myc-tagged protein but binding to AP-hNogo-A(1055-1120) was not observed under the conditions employed (Fig. 8).

Example 16: Identification of partial human NgR3 cDNA and protein sequences

The tblastn program was used to search for the human homolog of mouse NgR3. The mouse NgR3 protein sequence (SEQ ID NO:4) was used to query a proprietary human expressed sequence tag (EST) database from Incyte yielding one highly significant hit: Incyte Template ID 190989.1. This sequence (937 nucleotides)

- 135 -

contains an open reading frame of 312 amino acids in the second reverse frame that exhibits 88% identity with residues 66 to 381 of mouse NgR3 (SEQ ID NO:4), strongly indicating that it is part of the human NgR3 homolog.

5 A query of SEQ ID NO:4 against the public human EST database in Genbank also produced a hit with a 465-bp EST (Accession number: R35699; Version number: R35699.1; GI: 792600). There are a number of single nucleotide deletions and insertions within this sequence which cause frame shift errors. All of the reliable sequence contained in this public EST is present in the Incyte EST (Template ID 190989.1).

10 To obtain more nucleotide sequence that would extend the amino acid sequence at that carboxy terminal end, the I.M.A.G.E. Consortium clone No. 38319, which corresponds to Genbank accession No. R35699, was purchased from Incyte Genomics Inc. and subjected to further DNA sequence analysis. This clone consists of a NotI/HinD III fragment containing the sequence of interest, cloned into the
15 NotI/HinD III sites of the vector Lafmid BA (<http://image.llnl.gov/image/html/libs/lafmidBA.shtml>). The clone was received as an agar stab, which was streaked out on LB agar plates containing 50ug/ml ampicillin to isolate individual colonies. Six colonies were grown in LB medium with antibiotic, and plasmid DNA was prepared using the Promega Wizard Plus Miniprep DNA
20 Purification System (Promega #A7500). These DNAs were subsequently digested with NotI and HinD III restriction enzymes to confirm that the clones contained an insert. The insert of one isolate was sequenced using a combination of vector specific and gene specific primers yielding a partial nucleotide sequence of human NgR3 of 1176 nucleotides (SEQ ID NO:13). A translation of this sequence provides a partial
25 sequence for human NgR3 of 392 amino acids (SEQ ID NO:14).

The nucleotide sequence of SEQ ID NO:13 differs from the Incyte EST sequence at three positions. Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in the Incyte Template ID 190989.1 are GT (i.e., positions 12-13 of the complement of Incyte Template ID 190989.1). In
30 addition, position 641 in SEQ ID NO:13 is a C, whereas the corresponding nucleotide in the Incyte Template ID 190989.1 sequence is an A (i.e., position 641 of the complement of Incyte Template ID 190989.1). This results in two changes in amino

- 136 -

acids when comparing SEQ ID NO:14 to the ORF encoded by Incyte Template 190989.1: SEQ ID NO:14 contains a valine at position 5, whereas the ORF encoded by Incyte Template ID 190989.1 contains a leucine; SEQ ID NO:14 contains an alanine at position 214, whereas the ORF encoded by Incyte Template ID 190989.1 contains a glutamic acid.

The nucleotide sequence of SEQ ID NO:13 differs from the public EST (Accession number: R35699; Version number: R35699.1; GI: 792600) sequence at two positions (within the first 200 nucleotides of reliable sequence). Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in the public EST are GT (i.e., positions 12-13 of the public EST; Accession no: R35699; Version no: R35699.1; GI: 792600) This leads to a single amino acid change when comparing SEQ ID NO:14 to the ORF encoded by the public EST: SEQ ID NO:14 contains a valine at position 5, while the ORF encoded by the public EST contains a leucine.

A Bestfit analysis of the partial human amino acid sequence with the full-length mouse amino acid sequence indicates that the human NgR3 amino acid sequence is complete at the carboxy terminal end and that they share 89.54% identity. An alignment of all the NgR proteins is shown in Figure 9. Although the human NgR3 amino acid sequence is missing the first 25 amino acids, it can be determined that the human NgR3 protein contains the following features in common with the other NgR sequences: (1) eight Leucine Rich Repeat (LRR) domains; (2) an LRR carboxy-terminal (LRR-CT) domain; (3) a conserved cysteine in the fourth LRR domain; (4) a conserved potential glycosylation site in the eighth LRR domain; and (5) a hydrophobic carboxyl terminus.

As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by reference. This application claims benefit from United States provisional application 60/238,361, filed October 6, 2000, which is incorporated by reference herein in its entirety.

- 137 -

Key for Sequence Listing:

	SEQ ID NO:1	human NgR2 cDNA sequence derived from genomic sequence AC013606
	SEQ ID NO:2	human NgR2 amino acid sequence
5	SEQ ID NO:3	mouse NgR3 cDNA sequence derived from AC021768
	SEQ ID NO:4	a mouse NgR3 amino acid sequence
	SEQ ID NO:5	a human NgR1 amino acid sequence
	SEQ ID NO:6	a consensus amino acid sequence for NgRs
	SEQ ID NO:7	#1055-1120 amino acid residues of hNogoA (Nogo-66)
10	SEQ ID NO:8	a mature human NgR2 amino acid sequence
	SEQ ID NO:9	a mature mouse NgR3 amino acid sequence
	SEQ ID NO:10	a consensus NgR LLRNT amino acid sequence
	SEQ ID NO:11	a consensus NgR LRRCT domain amino acid sequence
	SEQ ID NO:12	a consensus NgR LRR domain amino acid sequence
15	SEQ ID NO:13	a partial human NgR3 nucleotide sequence
	SEQ ID NO:14	a partial human NgR3 amino acid sequence
	SEQ ID NO:15	a genomic sequence encoding a human NgR2 sequence.
	SEQ ID NO:16	a genomic sequence (complementary strand) encoding a mouse NgR3
20	SEQ ID NO:17	a mouse NgR1 amino acid sequence
	SEQ ID NO:18	a consensus sequence for the NTLRRCT domain of NgR
	SEQ ID NO:19	an consensus NgR LRRCT domain amino acid sequence

25

30

CLAIMS

What is claimed is:

5

1. An isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide comprising an LRRCT domain consisting of the amino acid sequence:

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

10

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

15

wherein X is any amino acid or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

20

2. The isolated nucleic acid according to claim 1, wherein X₁₇ and X₂₃ are independently selected from the group consisting of: arginine and lysine.

25

3. The isolated nucleic acid according to claim 2, wherein the amino acid sequence of the LRRCT domain is selected from the group consisting of: residues #261-310 of SEQ ID NO:2 and residues 261-310 of SEQ ID NO: 2 with up to 10 conservative amino acid substitutions.

30

4. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 2.

5. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 4 (mouse NgR3) or SEQ ID NO: 14 (human NgR3).

6. The isolated nucleic acid according to claim 1, wherein the

- 139 -

polypeptide comprises: (a) a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

5 7. The isolated nucleic acid to claim 1, wherein the polypeptide does not comprise an intact GPI domain.

8. An isolated nucleic acid consisting essentially of a nucleotide sequence complementary to a nucleotide sequence encoding a polypeptide selected
10 from the group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a polypeptide consisting of residues 256-396 of SEQ ID NO: 14 and a polypeptide consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to 100 nucleotides in length.

15 9. A vector comprising the nucleic acid of any one of claims 1, 4 or 5.

10. A host cell comprising a vector according to claim 9.

11. A polypeptide comprising a LRRCT amino acid sequence:

20

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

25 X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any amino acid residue or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

30

12. The polypeptide according to claim 11, wherein X₁₇ and X₂₃ is selected from the group consisting of arginine and lysine.

- 140 -

13. The polypeptide according to claim 11, wherein X₁₉ is glycine.

[SEQ ID NO:11]

14. The polypeptide according to claim 11, wherein the amino acid
5 sequence is selected from the group consisting of residues 261–310 of SEQ ID NO:2, residues 206–255 of SEQ ID NO: 14, residues 271–320 of SEQ ID NO:4 and amino acid sequences thereof comprising a conservative substitution.

15. A polypeptide comprising a NTLRRCT amino acid sequence:

10

C P X₁ X₂ C X₃ C Y X₄ X₅ P X₆ X₇ T X₈ S C X₉ X₁₀ X₁₁ X₁₂ X₁₃ X₁₄ X₁₅ X₁₆ P
X₁₇ X₁₈ X₁₉ P X₂₀ X₂₁ X₂₂ X₂₃ R X₂₄ F L X₂₅ X₂₆ N X₂₇ I X₂₈ X₂₉ X₃₀ X₃₁ X₃₂ X₃₃
X₃₄ F X₃₅ X₃₆ X₃₇ X₃₈ X₃₉ X₄₀ X₄₁ X₄₂ L W X₄₃ X₄₄ S N X₄₅ X₄₆ X₄₇ X₄₈ I X₄₉
X₅₀ X₅₁ X₅₂ F X₅₃ X₅₄ X₅₅ X₅₆ X₅₇ L E X₅₈ L D L X₅₉ D N X₆₀ X₆₁ L X₆₂ X₆₃ X₆₄
15 X₆₅ P X₆₆ T F X₆₇ G L X₆₈ X₆₉ L X₇₀ X₇₁ L X₇₂ L X₇₃ X₇₄ C X₇₅ L X₇₆ X₇₇ L X₇₈
X₇₉ X₈₀ X₈₁ F X₈₂ G L X₈₃ X₈₄ L Q Y L Y L Q X₈₅ N X₈₆ X₈₇ X₈₈ X₈₉ L X₉₀ D
X₉₁ X₉₂ F X₉₃ D L X₉₄ N L X₉₅ H L F L H G N X₉₆ X₉₇ X₉₈ X₉₉ X₁₀₀ X₁₀₁ X₁₀₂
X₁₀₃ X₁₀₄ F R G L X₁₀₅ X₁₀₆ L D R L L L H X₁₀₇ N X₁₀₈ X₁₀₉ X₁₁₀ X₁₁₁ V H X₁₁₂
X₁₁₃ A F X₁₁₄ X₁₁₅ L X₁₁₆ R L X₁₁₇ X₁₁₈ L X₁₁₉ L F X₁₂₀ N X₁₂₁ L X₁₂₂ X₁₂₃ L
20 X₁₂₄ X₁₂₅ X₁₂₆ X₁₂₇ L X₁₂₈ X₁₂₉ L X₁₃₀ X₁₃₁ L X₁₃₂ X₁₃₃ L R L N X₁₃₄ N X₁₃₅ W
X₁₃₆ C X₁₃₇ C R X₁₃₈ R X₁₃₉ L W X₁₄₀ W X₁₄₁ X₁₄₂ X₁₄₃ X₁₄₄ R X₁₄₅ S S S X₁₄₆
V X₁₄₇ C X₁₄₈ X₁₄₉ P X₁₅₀ X₁₅₁ X₁₅₂ X₁₅₃ X₁₅₄ X₁₅₅ D L X₁₅₆ X₁₅₇ L X₁₅₈ X₁₅₉ X₁₆₀
D X₁₆₁ X₁₆₂ X₁₆₃ C [SEQ ID NO:18]

20

25

wherein X is any amino acid residue or a gap and wherein the polypeptide is
not the polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17
(mouse NgR1).

16. The polypeptide according to claim 15, wherein X₆, X₃₇ and X₃₈
30 represents a gap.

17. A polypeptide comprising an amino sequence selected from the

- 141 -

group consisting of: SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:14.

18. The polypeptide according any one of claims 11, 15 or 17, wherein the polypeptide comprises: (a) an NTLRRCT domain, and (b) less than a complete
5 CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

19. The polypeptide according to any one of claims 11, 15 or 17, wherein the polypeptide does not comprise an intact GPI domain.

10

20. The polypeptide according to any one of claims 11, 15 or 17, wherein the amino acid sequence of the polypeptide further comprises an amino acid sequence of a heterologous polypeptide.

21. The polypeptide according to claim 20, wherein the heterologous polypeptide is an Fc portion of an antibody.

15

22. A method of producing a polypeptide according to any one of claims 11, 15 or 17, comprising the steps of introducing an isolated nucleic acid
20 according to any one of claims 1, 4, 5 or 8 or a vector according to claim 9 into a host cell, culturing said host cell under conditions suitable for expression of said polypeptide, and recovering said polypeptide.

23. An antibody that binds to a polypeptide of any one of claims 11, 15
25 or 17.

24. A composition comprising the polypeptide of claim 11, 15 or 17 and a pharmaceutically acceptable carrier.

25. A composition comprising the antibody of claim 23 and a pharmaceutically acceptable carrier.

30

- 142 -

26. A method of decreasing inhibition of axonal growth of a CNS neuron, comprising the step of contacting the neuron with an effective amount of the polypeptide of claim 11, 15 or 17.

5 27. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the polypeptide of claim 11, 15 or 17.

10 28. A method of decreasing inhibition of axonal growth of a CNS neuron comprising the step of contacting the neuron with an effective amount of the antibody according to claim 23.

15 29. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the antibody according to claim 23.

30. A method for identifying a molecule that binds a polypeptide of claim 11, 15 or 17 comprising the steps of:

- 20 (a) providing a polypeptide of claim 11, 15 or 17;
 (b) contacting the polypeptide with the candidate molecule;
 and
 (c) detecting binding of the candidate molecule to the polypeptide.

25

FIG. 1A

FIG. 1A

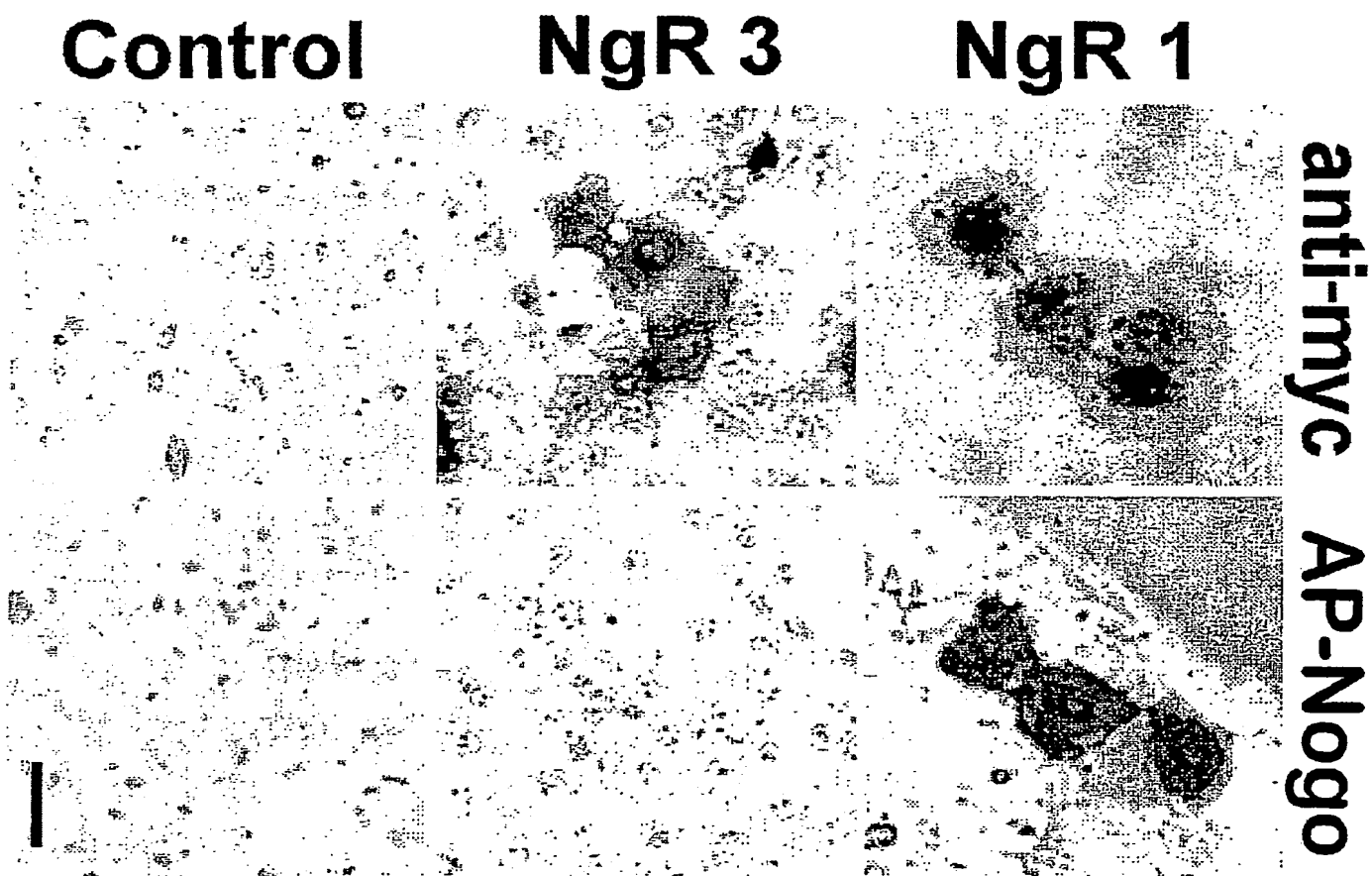
NOG0-R2	1	~~~~~MLPG	LRRLQAPAS	AC...LLLML	LA..LPLAAP	SCPMLCTCYS	50
NOG0-R3		MSWQSGTTVT	QSPVQAAQVS	CCCVELLLLL	LAGELPLGG	GCPRDCVCYP	
NOG0-R1		~~~~~MKRAS	AGGSRLLA	WV	LWLQAWQVAA	PCPGACVCYN	
Consensus		~~~~~S	-----S	-----LL---	L-----	-CP--C-CY-	
NOG0-R2	51	SP.PTVSCQA	NNFSSVPLSL	PPSTQRLFLQ	NNLIRTLRPG	TFGS...NLLT	100
NOG0-R3		AP.MTVSCQA	HNFAAIPEGI	PEDSERIFLQ	NNRITFLQQG	HFSP...AMVT	
NOG0-R1		EPKVTTSCPQ	QGLQAVPVG	PAASQRIFLH	GNRISHVPAA	SFRACRNLT	
Consensus		-P--T-SC--	-----P---	P-----R-FL-	-N-I-----	-F-----	
NOG0-R2	101	LWLFSSNNLST	IYPGTFRHLQ	ALEELDLCGN	RHLRSLEPDT	FQGLERLQSL	150
NOG0-R3		LWIYSNNITF	IAPNTFEGFV	HLEELDLCGN	RQLRTLAPET	FQGLVKLHAL	
NOG0-R1		LWLHSNVLAR	IDAAAFGLA	LLEQLDSDN	AQLRSVDPAT	FHGLGRLHTL	
Consensus		LW--SN----	I-----F----	-LE-LDL-DN	--LR---P-T	F-GL--L--L	
NOG0-R2	151	HLYRCQLSSL	PGNIFRGLVS	LQYLYLQENS	LLHLQDDLFA	DLANLSHLFL	200
NOG0-R3		YLYKCGLSAL	PAGIFGGLHS	LQYLYLQDNH	IEYLQDDIFV	DLVNLHLFL	
NOG0-R1		HLDRCGLQEL	CPGLFRGLAA	LQYLYLQDNA	LQALPDDTFR	DLGNLTHLFL	
Consensus		-L--C-L--L	-----F-GL--	LQYLYLQ-N-	---L-DD-F-	DL-NL-HLFL	
NOG0-R2	201	HGNRLRLTE	HVFRGLGSLD	RLLLHGNRLQ	GVHRAAFRGL	SRLTILYLFN	250
NOG0-R3		HGNKLWSLQ	GIFRGLVNLD	RLLLHENQLQ	WVHHKAFHDL	HRLTTLFLFN	
NOG0-R1		HGNRISSVPE	RAFRGLHSLD	RLLLHQNRA	HVHPHAFRDL	GRLMTLYLFA	
Consensus		HGN-----	--FRGL--LD	RLLLH-N---	-VH--AF--L	-RL--L-LF-	

FIG. 1B

Fig. 1B

NOC0-R2	251	NSLASLPGEA	LADLPSLEFL	RLNANPWACD	CRARPLWAWF	QRARVSSSDV	300
NOC0-R3		NSLTELGQDC	LAPLVALEFL	RLNGNAWDCG	CRARSLWEWL	RRFRGSSSAV	
NOC0-R1		NNLSALPTEA	LAPLRALQYL	RLNDNPWVCD	CRARPLWAWL	QKFRGSSSEV	
Consensus		N-L--L----	LA-L--L--L	RLN-N-W-C-	CRAR-LW-W-	---R-SSS-V	
NOC0-R2	301	TCATPPERQG	RDLRALREAD	FOAC...P.P	AAPTRPGSRA	350
NOC0-R3		PCATPELRQG	QDLKLLRVED	FRNC...TCP	VSPHQIKSHT	
NOC0-R1		PCSLPQRLAG	RDLKRLAAND	LQGCAVATGP	YHPITWTGRAT	DEEPLGLPKC	
Consensus		-C--P-----G	-DL--L---D	---C-----P	--P-----	-----	
NOC0-R2	351RGN	SSSNH.LY	G.VAE.....	AGAPPADPS.	...TLYRDLPA	400
NOC0-R3		.LTTSDRAA	..RKEHHP SH	G.ASRDKGHP	HGHPPGSRSG	YKKAGKNCTS	
NOC0-R1		CQPDAAADKAS	VLEPGRPASA	GNALKGRVPP	GDSPPGNGCSG	PRHI.NDSPF	
Consensus		-----	-----G	-----	--PP---S-	-----	
NOC0-R2	401	EDSRGR.....	QGGDAPTE.D	DYWGGY.....GGED	QRGEQMCPGA	450
NOC0-R3		HRNRNQISKV	SSGKELTELQ	DYAPDYQHKF	SFDIMPTARP	KRKGKCART	
NOC0-R1		GTLPCSAEPP	LTAVRPEGSE	P..PGFPTSG	PRRRPGCSRK	NRTRSHCRLG	
Consensus		-----	-----	-----	-----	-R-----	
NOC0-R2	451	ACQAPPDSRG	PALSAGLPSP	LLCLLLLVPH	HL~~~~~	~	491
NOC0-R3		PIRAPSGVQQ	ASSGTALGAP	LLAWILGLAV	TLR~~~~~	~	
NOC0-R1		QAGSGGGCTG	DSECGALPS	LTCSLTPLGL	ALVLWTVLGP	C	
Consensus		-----	-----	L-----	-L~~~~~	~	

3/6
FIG. 2



	1	50
Human NOGO-R1	~~~~~	GGSRLLAWVL WLQAWQAAP CPGACVCYNE
Murine NOGO-R1	~~~~~	~~~~~MKRASA GGSRLLAWVL WLQAWQAAP CPGACVCYNE
Murine NOGO-R3	~~~~~	~~~~~MKRASS GGSRLLAWVL WLQAWVATP CPGACVCYNE
Human NOGO-R3	~~~~~	MSWQSGTTVT QSPVQAAQVS GCCVELLLLL LAGELPIGGG CPRDCVCYPA
Human NOGO-R2	~~~~~	~~~~~MLPGLRRLQ APASACLLLM LLALPLAAPS CPMLCTCYSS
Consensus	~~~~~	~~~~~CP--C-CY--
	LRR NT	LRR 1
Human NOGO-R1	PKVTTSCPQQ	AASQRIFLHG NRISHVPAAS FRACRNLTIL
Murine NOGO-R1	PKVTTSCPQQ	ASSQRIFLHG NRISHVPAAS FQSCRNLTL
Murine NOGO-R3	P.MTVSCQAH	EDSERIFLQN NRITFLQOQH FSP..AMVTL
Human NOGO-R3	~~~~~	VDSERVFLQN NRIGLLQPGH FSP..AMVTL
Human NOGO-R2	P.PTVSCQAN	PSTQRLFLQN NLIRTLRPGT FGS..NLLTL
Consensus	P--T-SC--	----R-FL-- N-I-----F-----L
	LRR 2	LRR 3
Human NOGO-R1	WLHSNVLARI	LEQLDLSDNA QLRSDVPATF HGLGRHLTLH
Murine NOGO-R1	WLHSNALARI	LEQLDLSDNA QLVHVDPTTF HGLGHLTLH
Murine NOGO-R3	WIYSNNITFI	LEELDLDGDNR QLRTIAPETF QGLVKLHALY
Human NOGO-R3	WIYSNNITYI	LEELDLDGDNR QLRTIAPETF QGLVKLHALY
Human NOGO-R2	WLFSNNLSTI	LEELDLDGDNR HLRSLEPDTF QGLERLQSLH
Consensus	W--SN----I	LE-LDL-DN- -L-----P-TF -GL--L--L--L-
	LRR 4	LRR 5
Human NOGO-R1	LDRCGLQELG	QYLYQDNAL QALPDDTFRD LGNLTHLFLH
Murine NOGO-R1	LDRCGLRELG	QYLYQDNAL QALPDNTFRD LGNLTHLFLH
Murine NOGO-R3	LYKCGLSALP	QYLYQDNHI EYLQDDIFVD LVNLSHLFLH
Human NOGO-R3	LYKCGLSALP	QYLYQDNHI EYLQDDIFVD LVNLSHLFLH
Human NOGO-R2	LYRCQLSSLP	QYLYQENSL LHLQDDLFAD LANLSHLFLH
Consensus	L--C-L--L-	QYLYLQ-N-- --L-D--F-D L-NL-HLFLH

FIG. 3, cont.

	201	LRR 6	LRR 7	250
Human NOGO-R1	GNRISSVPER	AFRGLHSLDR	LLLHQNRVAH	VHPHAFRDLG
Murine NOGO-R1	GNRIPSVPEH	AFRGLHSLDR	LLLHQNHVAR	VHPHAFRDLG
Murine NOGO-R3	GNKLWSLGQG	IFRGLVNLDL	LLLHENQLQW	VHHKAFHDLH
Human NOGO-R3	GNKLWSLPGG	TFRGLVNLDL	LLLHENQLQW	VHHKAFHDLR
Human NOGO-R2	GNRLRLTTEH	VFRGLGSLDR	LLLHGNRLQG	VHRAAFRGLS
Consensus	GN-----	-FRGL--LDR	LLLH-N----	VH--AF--L-
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLYLFAN
				RLMTLYLFAN
				RLTTTLFLFNN
				RLTTTLFLFNN
				RLTILYLFNN
				RL--L-LF-N
				RLMTLY

FIG. 3, cont.

Human	NOGO-R1	401	LPGSAEPPLT	AVRPEGSEPP	GF...PTSGP	RRRPGCSRKN	RTRSHCRLGQ	450
Murine	NOGO-R1		LPSSAEPPLT	ALRPGGSEPP	GL...PTTGP	RRRPGCSRKN	RTRSHCRLGQ	
Murine	NOGO-R3		.SGKELTELQ	DYAPDYQHKF	SFDIMPTARP	KRKGKCARRT	PIRAPSGVQQ	
Human	NOGO-R3		.AGKQAPELP	DYAPDYQHKF	SFDIMPTARP	KRKGKCARRT	PIRAPSGVQQ	
Human	NOGO-R2		.QGGDAPTED	DYWGGY...GGED	QRGEQMCPGA	ACQAPPD...	
Consensus			-----	-----	-----	-R-----	-----	

Human	NOGO-R1	451	AGSGGGGTGD	SEGS	GALPSL	TCSLTPLGLA	LVLWTVLGPC	490
Murine	NOGO-R1		AGSGASGTGD	AECS	GALPAL	ACSLAPLGLA	LVLWTVLGPC	
Murine	NOGO-R3		ASSGTALG	APLLAWILGL	AVTLR~~~~	
Human	NOGO-R3		ASSASSLG	ASLLAWTLGL	AVTLR~~~~	
Human	NOGO-R2		SRGPALSA	GLPS PLLCL	LLLVPHHL~~	
Consensus			-----	-----	-----	-----	-----	

SEQUENCE LISTING

<110> BIOGEN, INC.
YALE UNIVERSITY
STRITTMATTER, STEPHEN M.
CATE, RICHARD L.
SAH, DINAH W.Y.

<120> NOGO RECEPTOR HOMOLOGS

<130> A116PCT

<140>

<141>

<150> 60/238,361

<151> 2000-10-06

<160> 16

<170> PatentIn Ver. 2.1

<210> 17

<211> 1260

<212> DNA

<213> Homo sapiens

<400> 1

```

atgctgcccc ggctcaggcg cctgctgcaa gctccccgct cggcctgect cctgctgatg 60
ctcctggccc tgcccctggc ggccccccagc tgccccatgc tctgcacctg ctactcatcc 120
ccgcccaccg tgagctgcca ggccaacaac ttctcctctg tgccgctgtc cctgccaccc 180
agcactcagc gactcttcct gcagaacaac ctcacccgca cgctgcggcc aggcaccttt 240
gggtccaacc tgctcaccct gtggctcttc tccaacaacc tctccaccat ctaccggggc 300
actttccgcc acttgcaagc cctggaggag ctggacctcg gtgacaaccg gcacctgcgc 360
tcgctggagc ccgacacctt ccagggcctg gagcggtgc agtcgctgca tttgtaccgc 420
tgccagctca gcagcctgcc cggcaacatc ttccgaggcc tggtcagcct gcagtacctc 480
tacctccagg agaacagcct gctccaccta caggatgact tgctcgcgga cctggccaac 540
ctgagccacc tcttcctcca cggaaccgc ctgcggtgc tcacagagca cgtgtttcgc 600
ggcctgggca gcctggaccg gctgctgctg cacgggaacc ggctgcaggc cgtgcaccgc 660
gcggccttcc gcggcctcag ccgcctcacc atcctctacc tgttcaacaa cagcctggcc 720
tcgctgcccc gcgaggcgct cgcgcacctg ccctcgctcg agttcctgcg gctcaacgct 780
aaccctggg cgtgcgactg ccgcgcgcgg ccgctctggg cctggttcca gcgcgcgcgc 840
gtgtccagct ccgacgtgac ctgcgccacc cccccggagc gccagggcgc agacctgcgc 900
gcgctccgcg aggcgcactt ccaggcggtg ccgcccgcgg caccacgcgc gccgggcagc 960
cgcgcccgcg gcaacagctc ctccaaccac ctgtacgggg tggccgaggc cggggcgccc 1020
ccagccgatc cctccaccct ctaccgagat ctgcctgccg aagactcgcg ggggcgccag 1080
ggcggggacg cgcctactga ggacgactac tgggggggct acgggggtga ggaccagcga 1140
ggggagcaga tgtgccccgc cgctgcctgc caggcgcccc cggactcccc aggccctgcg 1200
ctctcggccg ggctccccag ccctctgctt tgctcctgc tcctggtgcc ccaccacctc 1260

```

<210> 2

<211> 420

<212> PRT

<213> Homo sapiens

<400> 2

```

Met Leu Pro Gly Leu Arg Arg Leu Leu Gln Ala Pro Ala Ser Ala Cys
  1                      5                      10                      15

```

```

Leu Leu Leu Met Leu Leu Ala Leu Pro Leu Ala Ala Pro Ser Cys Pro
          20                      25                      30

```

Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys Gln Ala
 35 40 45
 Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr Gln Arg
 50 55 60
 Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly Thr Phe
 65 70 75 80
 Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu Ser Thr
 85 90 95
 Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu Leu Asp
 100 105 110
 Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr Phe Gln
 115 120 125
 Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln Leu Ser
 130 135 140
 Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln Tyr Leu
 145 150 155 160
 Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu Phe Ala
 165 170 175
 Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg Leu Arg
 180 185 190
 Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp Arg Leu
 195 200 205
 Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala Phe Arg
 210 215 220
 Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser Leu Ala
 225 230 235 240
 Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu Phe Leu
 245 250 255
 Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg Pro Leu
 260 265 270
 Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val Thr Cys
 275 280 285
 Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu Arg Glu
 290 295 300
 Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro Gly Ser
 305 310 315 320
 Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val Ala Glu
 325 330 335
 Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp Leu Pro
 340 345 350
 Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr Glu Asp
 355 360 365
 Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu Gln Met

370 375 380

Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly Pro Ala
385 390 395 400

Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu Leu Val
405 410 415

Pro His His Leu
420

<210> 3
<211> 1383
<212> DNA
<213> Mus sp.

<400> 3

atgtcttggc	agtctggaac	cacagtgaca	caatctcccg	tgcaggctgc	tcagggtctca	60
gggtgctgtg	tggaattgct	gctgttgctg	ctcgtctggag	agctacctct	gggtgggtgg	120
tgtcctcgag	actgtgtgtg	ctaccctgcg	cccatgactg	tcagctgcca	ggcacacaac	180
tttgctgcca	tcccggaggg	catcccagag	gacagtgagc	gcatcttcct	gcagaacaat	240
cgcatacact	tcctccagca	gggccacttc	agccccgcca	tggtcaccct	ctggatctac	300
tccaacaaca	tcactttcat	tgctcccaac	accttcgagg	gctttgtgca	tctggaggag	360
ctagaccttg	gagacaaccg	acagctgcga	acgctggcac	ccgagacctt	ccaaggcctg	420
gtgaagcttc	acgccctcta	cctctataag	tgtggactga	gcgccctgcc	cgcaggcatc	480
tttggtggcc	tgcacagcct	gcagtatctc	tacttgagg	acaaccatat	cgagtacctc	540
caagatgaca	tctttgtgga	cctggtcaat	ctcagtcact	tgtttctcca	tggttaacaag	600
ctatggagcc	tgggccaagg	catcttccgg	ggcctgggtga	acctggaccg	gttgctgctg	660
catgagaacc	agctacagtg	ggttcaccac	aaggctttcc	atgacctcca	caggctaacc	720
accctctttc	tcttcaacaa	cagcctcact	gagctgcagg	gtgactgtct	ggccccctg	780
gtggccttgg	agttccttcg	cctcaatggg	aatgcttggg	actgtggctg	ccgggcacgt	840
tccctgtggg	aatggctgcg	aaggttccgt	ggctctagct	ctgctgtccc	ctgcgcgacc	900
cccagagctgc	ggcaaggcca	ggatctgaag	ctgctgaggg	tggaggactt	ccggaactgc	960
acaggaccag	tgtctcctca	ccagatcaag	tctcacacgc	ttaccacctc	tgacagggct	1020
gccccgaagg	agcaccatcc	gtcccattgg	gcctccaggg	acaaaggcca	cccacatggc	1080
catccgcctg	gctccagggtc	aggttacaag	aaggcaggca	agaactgcac	cagccacagg	1140
aaccggaacc	agatctctaa	ggtgagctct	gggaagagac	ttaccgaact	gcaggactat	1200
gccccgcact	atcagcacaa	gttcagcttt	gacatcatgc	ccaccgcacg	acccaagagg	1260
aagggcaagt	gtgctcgag	gacccccatc	cgtgccccca	gtgggggtgca	gcaggcatcc	1320
tcaggcacgg	cccttggggg	cccactcctg	gcctggatac	tggggctggc	agtcactctc	1380
cgc						1383

<210> 4
<211> 461
<212> PRT
<213> Mus sp.

<400> 4

Met	Ser	Trp	Gln	Ser	Gly	Thr	Thr	Val	Thr	Gln	Ser	Pro	Val	Gln	Ala
1				5					10					15	
Ala	Gln	Val	Ser	Gly	Cys	Cys	Val	Glu	Leu	Leu	Leu	Leu	Leu	Leu	Ala
			20					25					30		
Gly	Glu	Leu	Pro	Leu	Gly	Gly	Gly	Cys	Pro	Arg	Asp	Cys	Val	Cys	Tyr
		35					40					45			
Pro	Ala	Pro	Met	Thr	Val	Ser	Cys	Gln	Ala	His	Asn	Phe	Ala	Ala	Ile
	50					55					60				
Pro	Glu	Gly	Ile	Pro	Glu	Asp	Ser	Glu	Arg	Ile	Phe	Leu	Gln	Asn	Asn
65					70					75				80	

Arg Ile Thr Phe Leu Gln Gln Gly His Phe Ser Pro Ala Met Val Thr
 85 90 95
 Leu Trp Ile Tyr Ser Asn Asn Ile Thr Phe Ile Ala Pro Asn Thr Phe
 100 105 110
 Glu Gly Phe Val His Leu Glu Glu Leu Asp Leu Gly Asp Asn Arg Gln
 115 120 125
 Leu Arg Thr Leu Ala Pro Glu Thr Phe Gln Gly Leu Val Lys Leu His
 130 135 140
 Ala Leu Tyr Leu Tyr Lys Cys Gly Leu Ser Ala Leu Pro Ala Gly Ile
 145 150 155 160
 Phe Gly Gly Leu His Ser Leu Gln Tyr Leu Tyr Leu Gln Asp Asn His
 165 170 175
 Ile Glu Tyr Leu Gln Asp Asp Ile Phe Val Asp Leu Val Asn Leu Ser
 180 185 190
 His Leu Phe Leu His Gly Asn Lys Leu Trp Ser Leu Gly Gln Gly Ile
 195 200 205
 Phe Arg Gly Leu Val Asn Leu Asp Arg Leu Leu Leu His Glu Asn Gln
 210 215 220
 Leu Gln Trp Val His His Lys Ala Phe His Asp Leu His Arg Leu Thr
 225 230 235 240
 Thr Leu Phe Leu Phe Asn Asn Ser Leu Thr Glu Leu Gln Gly Asp Cys
 245 250 255
 Leu Ala Pro Leu Val Ala Leu Glu Phe Leu Arg Leu Asn Gly Asn Ala
 260 265 270
 Trp Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Arg Arg
 275 280 285
 Phe Arg Gly Ser Ser Ser Ala Val Pro Cys Ala Thr Pro Glu Leu Arg
 290 295 300
 Gln Gly Gln Asp Leu Lys Leu Leu Arg Val Glu Asp Phe Arg Asn Cys
 305 310 315 320
 Thr Gly Pro Val Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr
 325 330 335
 Ser Asp Arg Ala Ala Arg Lys Glu His His Pro Ser His Gly Ala Ser
 340 345 350
 Arg Asp Lys Gly His Pro His Gly His Pro Pro Gly Ser Arg Ser Gly
 355 360 365
 Tyr Lys Lys Ala Gly Lys Asn Cys Thr Ser His Arg Asn Arg Asn Gln
 370 375 380
 Ile Ser Lys Val Ser Ser Gly Lys Glu Leu Thr Glu Leu Gln Asp Tyr
 385 390 395 400
 Ala Pro Asp Tyr Gln His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala
 405 410 415
 Arg Pro Lys Arg Lys Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala

5

260					265					270					
Ala	Trp	Leu	Gln	Lys	Phe	Arg	Gly	Ser	Ser	Ser	Glu	Val	Pro	Cys	Ser
		275					280					285			
Leu	Pro	Gln	Arg	Leu	Ala	Gly	Arg	Asp	Leu	Lys	Arg	Leu	Ala	Ala	Asn
		290				295					300				
Asp	Leu	Gln	Gly	Cys	Ala	Val	Ala	Thr	Gly	Pro	Tyr	His	Pro	Ile	Trp
		305				310					315				320
Thr	Gly	Arg	Ala	Thr	Asp	Glu	Glu	Pro	Leu	Gly	Leu	Pro	Lys	Cys	Cys
				325					330					335	
Gln	Pro	Asp	Ala	Ala	Asp	Lys	Ala	Ser	Val	Leu	Glu	Pro	Gly	Arg	Pro
			340					345					350		
Ala	Ser	Ala	Gly	Asn	Ala	Leu	Lys	Gly	Arg	Val	Pro	Pro	Gly	Asp	Ser
		355					360						365		
Pro	Pro	Gly	Asn	Gly	Ser	Gly	Pro	Arg	His	Ile	Asn	Asp	Ser	Pro	Phe
		370				375					380				
Gly	Thr	Leu	Pro	Gly	Ser	Ala	Glu	Pro	Pro	Leu	Thr	Ala	Val	Arg	Pro
		385				390					395				400
Glu	Gly	Ser	Glu	Pro	Pro	Gly	Phe	Pro	Thr	Ser	Gly	Pro	Arg	Arg	Arg
				405					410					415	
Pro	Gly	Cys	Ser	Arg	Lys	Asn	Arg	Thr	Arg	Ser	His	Cys	Arg	Leu	Gly
			420					425					430		
Gln	Ala	Gly	Ser	Gly	Gly	Gly	Gly	Thr	Gly	Asp	Ser	Glu	Gly	Ser	Gly
		435					440					445			
Ala	Leu	Pro	Ser	Leu	Thr	Cys	Ser	Leu	Thr	Pro	Leu	Gly	Leu	Ala	Leu
		450				455					460				
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys							
		465				470									

<210> 6

<211> 440

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus sequence

<220>

<221> MOD_RES

<222> (3)..(4)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (6)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (9)..(10)

<223> Variable amino acid

<220>
<221> MOD_RES
<222> (12)..(13)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (15)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (18)..(25)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (27)..(29)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (31)..(34)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (36)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (39)..(40)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (42)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (44)..(50)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (52)..(59)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (62)..(63)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (66)..(69)
<223> Variable amino acid

<220>
<221> MOD_RES

<222> (71)..(74)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (76)..(80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (87)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)..(91)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (94)..(96)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (98)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (101)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(105)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (111)..(112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>

<221> MOD_RES
<222> (116̄)..(117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (119̄)..(122)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (124̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (127̄)..(128)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (136̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (138̄)..(141)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (143̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (146̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (148̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (151̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (154̄)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162̄)..(170)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (175̄)..(176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (184)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (186)..(189)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (192)..(193)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (196)..(197)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (199)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (202)..(203)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (205)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (208)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (210)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (212)..(213)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (215)..(218)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (221)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (223)..(224)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (226)..(227)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (232)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (234)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (236)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (238)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (243)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (246)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (248)..(251)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (253)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (257)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (259)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (261)..(262)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (264)..(267)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (269)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (272)..(273)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (275)..(277)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (279)..(281)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (283)..(287)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (289)..(290)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (292)..(328)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (330)..(341)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (344)..(346)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (348)..(399)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (401)..(428)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (431)..(439)
 <223> Variable amino acid

<400> 6
 Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
 1 5 10 15

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Pro Xaa Xaa
 20 25 30
 Xaa Xaa Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa
 35 40 45
 Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser
 50 55 60
 Asn Xaa Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa
 65 70 75 80
 Leu Glu Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa
 85 90 95
 Pro Xaa Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa
 100 105 110
 Cys Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa
 115 120 125
 Leu Gln Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Xaa Leu Xaa Asp
 130 135 140
 Asp Xaa Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly
 145 150 155 160
 Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa
 165 170 175
 Leu Asp Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa
 180 185 190
 Xaa Ala Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa
 195 200 205
 Asn Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa
 210 215 220
 Leu Xaa Xaa Leu Arg Leu Asn Xaa Asn Xaa Trp Xaa Cys Xaa Cys Arg
 225 230 235 240
 Ala Arg Xaa Leu Trp Xaa Trp Xaa Xaa Xaa Xaa Arg Xaa Ser Ser Ser
 245 250 255
 Xaa Val Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Gly Xaa Asp Leu Xaa
 260 265 270
 Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Pro
 275 280 285
 Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 290 295 300
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 305 310 315 320
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 325 330 335
 Xaa Xaa Xaa Xaa Xaa Pro Pro Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa
 340 345 350
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

355 360 365
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 370 375 380
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg
 385 390 395 400
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 405 410 415
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa
 420 425 430
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu
 435 440

<210> 7
 <211> 66
 <212> PRT
 <213> Homo sapiens

<400> 7
 Arg Ile Tyr Lys Gly Val Ile Gln Ala Ile Gln Lys Ser Asp Glu Gly
 1 5 10 15
 His Pro Phe Arg Ala Tyr Leu Glu Ser Glu Val Ala Ile Ser Glu Glu
 20 25 30
 Leu Val Gln Lys Tyr Ser Asn Ser Ala Leu Gly His Val Asn Cys Thr
 35 40 45
 Ile Lys Glu Leu Arg Arg Leu Phe Leu Val Asp Asp Leu Val Asp Ser
 50 55 60
 Leu Lys
 65

<210> 8
 <211> 390
 <212> PRT
 <213> Homo sapeins

<400> 8
 Cys Pro Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys
 1 5 10 15
 Gln Ala Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr
 20 25 30
 Gln Arg Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly
 35 40 45
 Thr Phe Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu
 50 55 60
 Ser Thr Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr
 85 90 95

Phe Gln Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln
 100 105 110
 Leu Ser Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu
 130 135 140
 Phe Ala Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg
 145 150 155 160
 Leu Arg Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp
 165 170 175
 Arg Leu Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala
 180 185 190
 Phe Arg Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser
 195 200 205
 Leu Ala Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg
 225 230 235 240
 Pro Leu Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val
 245 250 255
 Thr Cys Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu
 260 265 270
 Arg Glu Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro
 275 280 285
 Gly Ser Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val
 290 295 300
 Ala Glu Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp
 305 310 315 320
 Leu Pro Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr
 325 330 335
 Glu Asp Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu
 340 345 350
 Gln Met Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly
 355 360 365
 Pro Ala Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu
 370 375 380
 Leu Val Pro His His Leu
 385 390

<210> 9
 <211> 421
 <212> PRT
 <213> Mus sp.

<400> 9

Cys Pro Arg Asp Cys Val Cys Tyr Pro Ala Pro Met Thr Val Ser Cys
 1 5 10 15
 Gln Ala His Asn Phe Ala Ala Ile Pro Glu Gly Ile Pro Glu Asp Ser
 20 25 30
 Glu Arg Ile Phe Leu Gln Asn Asn Arg Ile Thr Phe Leu Gln Gln Gly
 35 40 45
 His Phe Ser Pro Ala Met Val Thr Leu Trp Ile Tyr Ser Asn Asn Ile
 50 55 60
 Thr Phe Ile Ala Pro Asn Thr Phe Glu Gly Phe Val His Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg Gln Leu Arg Thr Leu Ala Pro Glu Thr
 85 90 95
 Phe Gln Gly Leu Val Lys Leu His Ala Leu Tyr Leu Tyr Lys Cys Gly
 100 105 110
 Leu Ser Ala Leu Pro Ala Gly Ile Phe Gly Gly Leu His Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Asp Asn His Ile Glu Tyr Leu Gln Asp Asp Ile
 130 135 140
 Phe Val Asp Leu Val Asn Leu Ser His Leu Phe Leu His Gly Asn Lys
 145 150 155 160
 Leu Trp Ser Leu Gly Gln Gly Ile Phe Arg Gly Leu Val Asn Leu Asp
 165 170 175
 Arg Leu Leu Leu His Glu Asn Gln Leu Gln Trp Val His His Lys Ala
 180 185 190
 Phe His Asp Leu His Arg Leu Thr Thr Leu Phe Leu Phe Asn Asn Ser
 195 200 205
 Leu Thr Glu Leu Gln Gly Asp Cys Leu Ala Pro Leu Val Ala Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Gly Asn Ala Trp Asp Cys Gly Cys Arg Ala Arg
 225 230 235 240
 Ser Leu Trp Glu Trp Leu Arg Arg Phe Arg Gly Ser Ser Ser Ala Val
 245 250 255
 Pro Cys Ala Thr Pro Glu Leu Arg Gln Gly Gln Asp Leu Lys Leu Leu
 260 265 270
 Arg Val Glu Asp Phe Arg Asn Cys Thr Gly Pro Val Ser Pro His Gln
 275 280 285
 Ile Lys Ser His Thr Leu Thr Thr Ser Asp Arg Ala Ala Arg Lys Glu
 290 295 300
 His His Pro Ser His Gly Ala Ser Arg Asp Lys Gly His Pro His Gly
 305 310 315 320
 His Pro Pro Gly Ser Arg Ser Gly Tyr Lys Lys Ala Gly Lys Asn Cys
 325 330 335
 Thr Ser His Arg Asn Arg Asn Gln Ile Ser Lys Val Ser Ser Gly Lys

```
<210> 11
<211> 50
<212> PRT
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: Consensus
sequence

<220>
<221> MOD_RES
<222> (2)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (4)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (6)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (11)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (14)
<223> Variable amino acid.

<220>
<221> MOD_RES
<222> (16)..(19)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (21)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (25)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (27)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (29)..(30)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (40)..(41)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (43)..(45)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (47)..(49)
 <223> Variable amino acid

<400> 11
 Asn Xaa Trp Xaa Cys Xaa Cys Arg Ala Arg Xaa Leu Trp Xaa Trp Xaa
 1 5 10 15
 Xaa Xaa Xaa Arg Xaa Ser Ser Ser Xaa Val Xaa Cys Xaa Xaa Pro Xaa
 20 25 30
 Xaa Xaa Xaa Gly Xaa Asp Leu Xaa Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa
 35 40 45
 Xaa Cys
 50

<210> 12
 <211> 196
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Consensus
 sequence

<220>
 <221> MOD_RES
 <222> (2)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (5)..(6)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (8)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (10)..(16)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (18)..(25)
 <223> Variable amino acid

<220>

<221> MOD_RES
<222> (28)..(29)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)..(40)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (42)..(46)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (49)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (53)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (56)..(57)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (60)..(62)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (64)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (67)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (70)..(71)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (73)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (75)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (77)..(78)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (82)..(83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (85)..(88)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (93)..(94)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (102)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (120)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (128)..(136)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (141)..(142)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (150)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (152)..(155)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (158)..(159)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162)..(163)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (165)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (168)..(169)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (171)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (174)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (178)..(179)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (181)..(184)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (187)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (189)..(190)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (192)..(193)
 <223> Variable amino acid

<400> 12
 Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser Asn Xaa
 20 25 30
 Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Leu Glu
 35 40 45
 Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa Pro Xaa
 50 55 60
 Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa Cys Xaa
 65 70 75 80
 Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa Leu Gln
 85 90 95
 Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp Asp Xaa
 100 105 110
 Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly Asn Xaa
 115 120 125
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa Leu Asp
 130 135 140
 Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa Xaa Ala
 145 150 155 160
 Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa Asn Xaa
 165 170 175
 Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa Leu Xaa
 180 185 190
 Xaa Leu Arg Leu
 195

<210> 13
 <211> 1176
 <212> DNA
 <213> Homo sapiens

<400> 13
 gagggcatcc ccgtggacag cgagcgcgtc ttctgcaga acaaccgcat cggcctcctc 60
 cagcccggcc acttcagccc cgccatggtc accctgtgga tctactcgaa caacatcacc 120

```

tacatccacc ccagcacctt cgagggtctc gtgcacctgg aggagctgga cctcggcgac 180
aaccggcagc tgcggacgct ggcacccgag accttccagg gcctgggtgaa gcttcacgcc 240
ctctacctct acaagtgtgg gctcagcgcc ttgccggccg gcgtcttttg cggcctgcac 300
agcctgcagt acctctacct gcaggacaac cacatcgagt acctccagga cgacatcttc 360
gtggacctgg tcaacctcag ccacctgttt ctccacggca acaagctgtg gagtctgggc 420
ccgggcacct tccggggcct ggtgaacctg gacctctttt tgctgcacga gaaccagctg 480
cagtgggtcc accacaaggc attccacgac ctccgcaggc tgaccaccct cttcctcttc 540
aacaacagcc tctcggagct gcagggtgag tgcctggccc cgctgggggc cctggagttc 600
ctccgcctca acggcaaccc ctgggactgt ggttgctcgc gcgctccct gtgggaatgg 660
ctgcagaggt tccggggctc cagctccgct gtcccctgtg tgtcccctgg gctcggcac 720
ggccaggacc tgaagctgct gagggccgag gacttccgga actgcacggg accagcgtcc 780
ccgcaccaga tcaagtcaca cacgtcacc accaccgaca gggccgcccg caaggaacac 840
cactcaccac acggcccccac caggagcaag ggccaccgc acggcccccg gcccgccac 900
aggaagccgg ggaagaactg caccaacccc aggaaccgca atcagatctc taaggcgggc 960
gccgggaaac aggccccga gctgccagac tatgccccag actaccagca caagttcagt 1020
tttgacatca tgcctacggc ccggcccaag aggaagggca agtgtgcccg caggaccccc 1080
atccgtgccc ccagcggggt gcagcaggcc tctcggcca gttccctggg ggccctccctc 1140
ctggcctgga cactggggct ggcggtcact ctccgc 1176

```

<210> 14

<211> 392

<212> PRT

<213> Homo sapiens

<400> 14

```

Glu Gly Ile Pro Val Asp Ser Glu Arg Val Phe Leu Gln Asn Asn Arg
 1             5             10             15

```

```

Ile Gly Leu Leu Gln Pro Gly His Phe Ser Pro Ala Met Val Thr Leu
      20             25             30

```

```

Trp Ile Tyr Ser Asn Asn Ile Thr Tyr Ile His Pro Ser Thr Phe Glu
      35             40             45

```

```

Gly Phe Val His Leu Glu Glu Leu Asp Leu Gly Asp Asn Arg Gln Leu
      50             55             60

```

```

Arg Thr Leu Ala Pro Glu Thr Phe Gln Gly Leu Val Lys Leu His Ala
      65             70             75             80

```

```

Leu Tyr Leu Tyr Lys Cys Gly Leu Ser Ala Leu Pro Ala Gly Val Phe
      85             90             95

```

```

Gly Gly Leu His Ser Leu Gln Tyr Leu Tyr Leu Gln Asp Asn His Ile
      100            105            110

```

```

Glu Tyr Leu Gln Asp Asp Ile Phe Val Asp Leu Val Asn Leu Ser His
      115            120            125

```

```

Leu Phe Leu His Gly Asn Lys Leu Trp Ser Leu Gly Pro Gly Thr Phe
      130            135            140

```

```

Arg Gly Leu Val Asn Leu Asp Arg Leu Leu Leu His Glu Asn Gln Leu
      145            150            155            160

```

```

Gln Trp Val His His Lys Ala Phe His Asp Leu Arg Arg Leu Thr Thr
      165            170            175

```

```

Leu Phe Leu Phe Asn Asn Ser Leu Ser Glu Leu Gln Gly Glu Cys Leu
      180            185            190

```

```

Ala Pro Leu Gly Ala Leu Glu Phe Leu Arg Leu Asn Gly Asn Pro Trp
      195            200            205

```

Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Gln Arg Phe
 210 215 220
 Arg Gly Ser Ser Ser Ala Val Pro Cys Val Ser Pro Gly Leu Arg His
 225 230 235 240
 Gly Gln Asp Leu Lys Leu Leu Arg Ala Glu Asp Phe Arg Asn Cys Thr
 245 250 255
 Gly Pro Ala Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr Thr
 260 265 270
 Asp Arg Ala Ala Arg Lys Glu His His Ser Pro His Gly Pro Thr Arg
 275 280 285
 Ser Lys Gly His Pro His Gly Pro Arg Pro Gly His Arg Lys Pro Gly
 290 295 300
 Lys Asn Cys Thr Asn Pro Arg Asn Arg Asn Gln Ile Ser Lys Ala Gly
 305 310 315 320
 Ala Gly Lys Gln Ala Pro Glu Leu Pro Asp Tyr Ala Pro Asp Tyr Gln
 325 330 335
 His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala Arg Pro Lys Arg Lys
 340 345 350
 Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala Pro Ser Gly Val Gln
 355 360 365
 Gln Ala Ser Ser Ala Ser Ser Leu Gly Ala Ser Leu Leu Ala Trp Thr
 370 375 380
 Leu Gly Leu Ala Val Thr Leu Arg
 385 390

<210> 15
 <211> 143899
 <212> DNA
 <213> Homo sapiens

<220>
 <221> modified_base
 <222> (2044)..(2144)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (6609)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (6625)..(6724)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (14153)..(14252)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base

<222> (19512)..(19611)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (22595)..(22694)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (27825)..(27924)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (34953)..(35052)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (40783)..(40882)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (49000)..(49099)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (62884)..(62983)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (75528)..(75627)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (87944)..(88043)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (111030)..(111129)
 <223> a, t, c, g, other or unknown

<400> 15
 aagcacatac aggtgacatt acagaactga cagttatgcc aggcactgta cttagcccct 60
 ataccatcct caaacagctg tatgatgtag attgggtatt aacccatta ataacaaaag 120
 tacagggaac aaagtgactt tccaaaggtc atgccattca aaggagggtg aatcttaggt 180
 tggacgcagg ctgtctgact ctggagtctg aggtgttaat gctgcctcct ccatgggaac 240
 agcccaagtg aaaaacagct gatccactct tcatttactt ggcatctgtg ctaagctggt 300
 ccctgagcca agctctgagc aacagaaaca gaagctctgc attaggagct tgtgagcatg 360
 tcaatgccgg gtaaaggagt gctggaaacc gctgggatgg ccgccgagca ctaggccgtt 420
 gaaggtgggc tctgtgtgac tgggttcctct acactctggc ctggctgcct gcaggaagaa 480
 gatcaagctg agtgggctgg ccctggacca caaggtgaca ggtgacctct tctacacca 540
 tgtgaccacc atgggccaga ggctcagcca gaaggccccc agcctggagg acggttcgga 600
 tgccttcattg tcaccccagg atgttcgggg cacctcagaa aaccttcctg agagtgagtg 660
 tctggtcaag gtgccggcct tgggggatag tgatgggtgg tcctcatatt cagtgagcac 720
 tcatggttga gtattttatt gcacccctct tcagtcctta caacacccca tgatgtaggt 780
 ggggcatgct cctcatttac agatgggcac atcaaagctc agctaacgtg gggaaagttca 840
 gattcagggt tacctgtctg gattcctggg attggggagg gaggagcttc caaatgggg 900

acaagggtctc	tgggcctgtc	gggtagctgg	tttcctcagg	gccccttgca	acctctgagc	960
ttattgcac	aggtgcagcc	aggcccgtga	gcctcctggc	aggggtcctc	cacacctggc	1020
tgtcttttgc	cccttgctgg	tcacaggagg	agctgcagca	cctgcctggg	ctgcttctca	1080
ggagggtaca	tgaagatccc	aggaccgcca	gctccatgat	aagtgggaagg	agctccttgg	1140
agtcaggagc	gggagttgag	gagtttgagt	cctgctctcc	agttataggc	tatgtgactt	1200
gtgtagatca	cctaacccttg	ctcttgattt	ccttacctct	taaactagca	ctaaaagcac	1260
cccacaaact	gtaaagttag	ttgtgatgat	tgaatgacac	catgggtgtg	gaagctcttt	1320
gtaaagtgc	aaacgggtgtg	cagtttgagg	gtgggttacc	ccagtgcga	ttctcagagg	1380
gcaacatggc	taagggcacg	agctggagtt	aggctgacct	gctgcttcca	gccctgtgag	1440
cttgagcaag	tcatttaact	tcctgagctg	cagtttcttc	atcagtaaaa	tgtaaaagg	1500
atagggtgtg	tgtgaagattt	tattaaatgg	ggtaataaat	gtcaagtatg	tagcccatag	1560
tgagtgtctc	agagtttttt	tcttttggtt	ctttccccc	cgccccgaga	tggagcctta	1620
ctctgttgcc	caggctggag	tgcaagtggc	tgatcttggc	tactgcaac	ctccgctcc	1680
cgggttcaag	caattctcct	gcctcagcct	cccaaatagc	tgggactaca	ggcgtgcacc	1740
accatgctcg	gctaattttt	gtatctttag	tagagacggg	gtttcaccat	gttggccagg	1800
ctggtctcga	actcctgacc	tcagtatgct	cctgcctcag	ccccgaaa	ttttgggatt	1860
acaagtgtga	gccccgtgc	cctgcccaggt	tttttttttt	tttttttttt	tgtaaaacac	1920
ccacagggtg	ttgctgttgc	ctgggctgga	gtgctgtagt	gcaatcatag	ttcactgcag	1980
ccttgacctc	ctgggctcaa	gtgatcctcc	tgctcagcc	tcctgagtag	ctgggaatac	2040
aggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnttttgtg	ttttaagtag	2160
agacagggtt	ttcccaatgt	tggccaaggc	tgggtctaaa	ctcccaacct	cagggtgatcc	2220
accacctca	gcctcccaaa	gtactgggat	tacaggcgtg	agacaccgtg	cccagccagg	2280
aggctatttt	ctctgataaa	ttaccagctc	tcagggtatt	ctctacagcg	atgcaagaac	2340
agcctaatac	atccaggctc	agcatcagtg	gaccagggtg	ggagagctta	agatgtcaag	2400
gtctgaatgc	cgcttccaca	cacctttggg	acctaggggc	tcctctctct	tttctttttt	2460
cagtagaaga	tgttatcttc	tcctttctct	gaccagtagt	tggtgatggg	ttcagagata	2520
gtttttcagt	caagatatat	ttcagtggct	tcactgagcc	caagtccct	cgctctctta	2580
ggactttatt	tccttggttt	tagaagaggg	ataacacata	ttttctaagg	tggttgtgag	2640
attaagggag	ctgggtaccg	gtggtgcata	aggacaggat	agagcaatgg	tgagccact	2700
caaaaagcga	aaagttgacc	tgcgagggtg	acacttatca	aatcagcaca	cagtgggagt	2760
ggaaggaatg	tcctcatca	gttacaatat	ttggagagtg	caagttatag	aaaaccagc	2820
cctggccggg	cgcggtgggt	catgcctata	atcccagcac	tttgggaggc	tgaggcagg	2880
ggatcacgag	gtcaggaggt	caagaccagc	ctgaccaacg	tggtgaaacc	ccacctctac	2940
taaaaataca	aaattagctg	ggcgtgggtg	tgtgtgcctg	taactctgagc	tactcaggag	3000
gctgaggcac	gagaatcact	tgaaacggg	aggtggagtt	tgcaagtgagc	cgagatcgca	3060
ccactgcact	ccagcctggg	caacagagcg	agactccatc	tcaaacgaaa	aaaaaaaaag	3120
aaagaaaacc	cagctctaac	tggcttaaac	agtaagaaga	tctattatat	tatccatctc	3180
aggcagcagc	aagcccagag	gtaggggact	ccaagggttg	ttgatccagg	gcttaacgat	3240
gtcatcaaag	accaggttc	tttctgtctc	ggcacctctg	tctgcagggc	cagcttcatc	3300
ctaagccaga	ttgttcttgt	cttgattaca	agttggctgc	tgggccagca	gacgtgcct	3360
gcctccctgt	tcactttcag	aagtagaaag	tggcccttcc	ccagtcatgg	aatgaaagag	3420
tttctttctt	gtctgggatt	gcttaggtcc	acccacctga	agccaatgac	tgtaaccagg	3480
aagctaatat	acactgattg	tcttaagtca	gggttctctg	gccagtcttg	gtcaaggagt	3540
gtgatactgt	catgattgtc	ttgggctcat	cagggcagct	ctgcagatga	gatcaaaactc	3600
caagctacat	tattctgaac	agtgggaagt	aggaaagaga	cattttggga	gatacaaaac	3660
acaatgtcta	tcccatatcc	ctaggtccag	gtcacagtgt	cttggttgga	catcaaatgt	3720
agaaaaagaa	agactgtcca	tccatttatc	tacctattca	tctggttttt	gatttttttt	3780
aaattttatt	ttaagacatt	ctcactctgt	caaccagact	ggagtgcagt	ggtttgatca	3840
tggtcatgag	cagcctcaac	ctcccaggct	caagtgaacc	tcccatgctc	aagtgtacct	3900
cctacctcag	cctcccaagt	agctagaact	aaaggtgcat	gccaccacgc	tcagttaatt	3960
tttgcatttt	ttgtagagat	ggggtttcgt	catgatgccc	atgctagtct	ggaattcctg	4020
aactcaagca	atatgcctgc	ctttgcctcc	caaatgctg	ggattgtagg	catgagccac	4080
tgtccttgcc	tcatctgttt	aataatttat	gaaacaacta	ctgggtgctg	agcacggggc	4140
caggggctgg	agatctagca	gggaccaggc	agatctctgc	caagtctgtg	gtttcttaaa	4200
ggttttgtct	ataattcccc	ttttcttttc	tctttcgttt	ttttcttttt	ctttctttct	4260
ttctttcttt	tttttttttt	gagacagagt	ctcactctgt	taccaggct	ggagtgcagt	4320
ggtgcgatct	cagctcactg	caacctctgc	ctcctgggtt	caagcgattc	tcctgcctca	4380
gcctcccag	tagctgggac	tacaggcgcc	tgccaccatg	cccggctaatt	ttttgtgttt	4440
ttagtagaga	ctgggtttca	ccatattggc	caggctggtc	ttgaactcct	gacctgtgga	4500
tccgcccgt	tcggcctccc	acagtgtgg	gattacaggc	gtgagccacg	gcgcccagcc	4560
agtttccctt	ttcaatgagg	cctccctgac	ctccatactc	tactcctcca	cctggcccac	4620
tcagctctac	cttttcttcc	ccatagcact	caagacctcc	taacatacta	cgtaagttat	4680
ttatttacta	ggcttactgt	gtattgtctg	tcttcctcta	ctagaatgta	aactccatga	4740

gaatagaaat	ttttgccttt	ttattttagt	tggtgtctgc	agccctggc	ttagtccctg	4800
gcatacaaca	gtcactccac	ccacagttgc	tgaataagtg	actaaaggtc	cctgcccctca	4860
tattgttatg	agggagtgtg	catgtttgta	gagaaaaatc	tgaggcacia	taaaattttta	4920
tagagtttaa	gttttctttt	ttaagcaatc	cacgaattgg	ggtagtttca	gaggtagttt	4980
ttcagtcattg	acgtatttca	atggcttcac	tgagcccaag	ttctttcacc	tctctaggac	5040
tttatttcct	tatttctaga	acggggataa	cacatagttc	ataaggcagt	tatgagagta	5100
agggagctgg	tatgggggtg	tgacataagga	caggatagag	cagtgggtgag	accgctcaga	5160
tgacaaaagcg	tcagagacca	gtattttacga	cggaaatgtg	gaagcatgat	aaagaaatta	5220
tttgggctgg	gcacaatgac	tcacaactaa	taaaactttg	ggaggccaag	gtgggaggat	5280
cacttgacatt	gcagaaggtc	aaggctgcag	tgagctgtga	ttttgpcact	gcactccagc	5340
ctgggtcaaca	gagtgagacc	ctggctcgaa	acgttatttg	attggttaca	gttatacagt	5400
tgccctatttt	ggtctattcc	atttgaaagt	tcctagttct	ataattttta	gtttgttggc	5460
tgtttctgat	tggttaagct	taagttttgt	tttcccttaa	tacagtttaag	tgccccataa	5520
tgacatttttg	gtcaaggaca	gaccacatat	acagtgggtg	tcccataaga	ttataatgga	5580
gctgaaacat	tcctattgtc	tatggcgtag	tggtcctgat	gttgtagcgc	aatgcattag	5640
ttatatgttt	gtggcaatgc	tggtgtaaac	acacactctg	cactgccagt	gatataaaaag	5700
aatagcacat	acagttatat	atagtacata	atatctgata	atgataatac	ataactatat	5760
tactggttta	tatatattact	atattattta	tctttatttt	atttttgaga	cagagtctca	5820
ttctgtcacc	caggctggag	tgacgtggcg	cgatcttggc	tcaccgcaac	ctccgcttcc	5880
tggtttcaag	tgattctcct	gcctcagctc	cctgagtagc	tggtgattaca	ggtgtgcacc	5940
atgacaccct	gctaatatgt	tttgattttt	tagtagagat	ggggtttcac	catgttggcc	6000
aggctggctc	tgaactactg	acctcaagtg	atcaccccg	ctcggttcc	caaagtctgt	6060
ggattacagg	cgtgagccac	cacgcatggc	ctattttata	ttattttaga	gtgtacgcct	6120
tatacttata	aaaaaaagct	aactgtcaaa	cagcctcggt	caggctcctc	aacagatatt	6180
ccagaagaca	ttgttatcat	aggagatgac	agctccgtgc	atattattgt	ccctgaaaac	6240
cttctagtgt	ggaagtggaa	gacagtgata	ttgatgatag	gacccagtgt	aggcctaggc	6300
taatgtgtgt	gtttgtgtct	ttgcttttaa	caagaaagtt	taaaaagtta	aaataaaata	6360
caaaaatttt	taaatagaaa	aaagctgccc	aggaacaatg	gctcacacct	gtaatcccac	6420
cattcgggga	ggccaagggtg	ggtggattgc	ttgagctcag	gagttcaaga	ccagcctggg	6480
caacatgggtg	aaaccccatc	tctacaaaaa	atacaaaaat	tagccgggtg	tggtggcatg	6540
cggctatagt	tccagctaatt	cgaggggctg	aggtgggagg	atcactgggg	gggagggtgt	6600
tgaggctgna	gtgagctgtg	attgnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6720
nnnnnatatc	ttaaaaaat	ttttttttat	ttttgagaca	gaattttctc	cttgttgccc	6780
aggctggagt	gcaatggcgc	tatctcagct	cagggaaccc	tccacctcct	gggttcaagc	6840
gattctcctg	ccttagcctc	ccaggtacag	gcgcccgcga	ccatgctcgg	ctaatttttg	6900
tattttttagt	agagatgggg	tttcaccatg	ttgtccaggc	tggtcttgaa	atcctgcctc	6960
aggatgatcca	ccccctcgg	cctcccaaag	tgctggaatt	tacaggcgtg	agccactgtg	7020
cctggcctcc	tttacatttt	tttaaattta	atttttaatt	tttaattttt	aattttctcat	7080
atatatatat	ttttaagact	agccaagtga	agcagtggga	gtgggaaaag	aactggtttt	7140
gatcaatagg	tgtaaacacc	actgcactgg	gaccagccta	ttttacattc	ctgttagcag	7200
tgatgagggg	tcactttctt	tgtagcctca	acaatatgtg	tcgttgccca	tcttttttct	7260
tttttttttt	tttttttttg	agatggagtc	tcactctgtt	gcctaggctg	gaatgcaatg	7320
gcatgatctc	agctcactgc	aacctccgcc	tcccaggttc	aagtgtattc	tgtgtctcag	7380
cctcctgagt	agatgggatt	acaggcgtcc	accaccacgc	ccggctaatt	ttttgtattt	7440
tcagtagaga	tggtgtttca	ccatgttggc	caggttgggt	tcgaactcct	gacctcaagt	7500
gatccgcccc	cctcggcctc	ccaaagtgtc	gggattacag	gcatgagcca	ccgcgccegg	7560
cctgcccctc	ttttttttgt	tatagccatc	ctagtggatg	taaagttttt	ttgtgatttt	7620
gatttgtgtt	tcctactga	tcaatgatgt	tgagcatcct	ttcctgtgct	tattggcctt	7680
tggtatatct	tgagagaaag	gtctattcag	gtcctttgcc	cacttttaaaa	ttaggttata	7740
tttctattac	tgagatgtaa	gagttcttta	tgttctagat	ataagtctcc	tacatatgat	7800
ttgtaaaaat	tttcctttca	ttattgggtt	gtctttcact	ttcttttggg	gtccttttagt	7860
gcacaacagt	ttttaaatatt	gaagtccaat	tttctatttt	tctcttttgc	cacttgtatc	7920
ttggtgtcat	gtttaaggaa	ctattgccta	atctcagggt	acaaagattt	acacctgtgt	7980
ttccttcttt	ccttctctcc	ttccttctct	ccttctctcc	ctccctccct	ccttctctcc	8040
ctccctctct	ccttctctcc	ctccctctcc	ctccctctcc	ctccctctcc	ctccctctcc	8100
ttccttctct	ccttctctcc	ctccctctcc	ccttctctcc	ctccctctcc	ctccctctcc	8160
tctgtcacc	aggtgggagt	gtagtggcac	gatcttgggt	cactgcaacc	tctgctcct	8220
gggttcaagc	aattctcctg	cctcagcctc	ctgagtagct	gggactacag	gcacacacca	8280
ccatgcccag	ctaatttttg	tatttttagt	agagacgggg	tttcaccaca	ttggccagga	8340
tggtttcgat	ctcctgacct	cgtgatccac	ccgccttggc	ctcccaaagt	gctgggattg	8400
caggtgtgag	ccaccatgcc	cggcctgtgt	tttcttagat	ttttgtagtt	ttagctctta	8460
tagttagatc	cttgatccat	tttgagttga	ttttgtatat	agtgtgagat	atccacctgg	8520
tggtgtaaat	tgcccagaag	tggtgtatgt	tctaaatctg	gctgttaggg	attactagag	8580

gtgaccaaag	tgaatTTTTT	ctttgtttct	tttttttttt	ggagacagag	tctccgtcac	8640
ccaggctgga	gtgcaatggc	ttcatcttgg	ctcagtgcaa	cctctgcctt	ctggtttcaa	8700
gcagttctcc	tgccctcagac	tccctgagtag	ctgggtattac	aggcgtgtac	caccatgctt	8760
ggctaatttt	tgtatTTTTa	gtaaagatgc	agtttcacct	gttggccagg	cttttctgga	8820
actccccgcc	tcaagtgatc	catctgcctc	tacctcccaa	agtgtctggga	ttacaggtgg	8880
gagccaccgt	gcccagtcct	tttctcagaa	tttatTTgtt	tttttttgtt	ttgtttcatt	8940
tttgagatag	ggtctcactc	tgctcagctag	gcaggagttc	agtgggtgta	tcattgtctgc	9000
agccttgaac	ttctggactc	acgtgatctt	cccacctcag	cctcctgagt	agctaggatt	9060
acaggcatgt	gcttccacac	ctggctaatt	ttttaatttt	ctaggactta	tttgtccatt	9120
cttgcaa>gc	agggtacaac	atgcctatct	ctacctacct	ctcttccctt	caagggactc	9180
cagccaaaat	ccttgaggct	ctcgggctga	ctgtgggtgc	tgttgcttga	tctgcctcag	9240
tcatgtctga	tgatcaaaaag	tgtccgtttt	ctgcttcttg	gaactttatt	cactttgggt	9300
gtcagtcctc	ctctgcagtg	tcccaagaac	acagaattag	accaggaatc	tgtgttgcca	9360
tagtgtgtgg	aaagaggcag	acttccaact	ccgctatgtg	ctgttggttg	attgaagctt	9420
aattttcttt	ctatctttct	ttcttttctt	ttcttttttt	tttttttgag	atggaactct	9480
gctctgttgc	ccaggctgga	gtgcagtggt	gcgatctcac	ctcactgcaa	cctccgcctc	9540
ccagggtcaa	gcgattctcc	tgccctcagcc	tccctgagtag	ctgggattac	aggtgcatgc	9600
caccatgccc	ggctaatttg	tgtaatttta	gtagaaacag	tgtttcacca	tattggctcag	9660
gctggctctg	acctcctcac	ctcaggtgat	ccacccgcct	tggcctccca	aagtgtcggg	9720
attacaggcg	tgagccaccg	tgccctggcac	ttaattttct	taatacctca	attaccccat	9780
atggtaaaaat	gggactagta	atccatacct	tatagcgctg	ttgtgaaaat	gaaatgaggg	9840
taagcagata	aaatttcaga	ctacggatgg	gattgttact	acattctgaa	cctggctttg	9900
ctgttatattg	ctatgtgacc	ttatcttctc	tggaatccca	ttctttccaa	gtctataaaa	9960
caaagtggag	aattgtcaac	ctttcttcca	aagagcaatg	atttaaggat	caaatgatgt	10020
catttaacaa	aaatatgaag	agctcaacaa	atgaggaact	cattattatt	attacaatta	10080
ttattatTTT	agaaataggg	tcttgttctc	ttgcctaggc	tggagtccag	tggtataaac	10140
acagctcaat	gcatcttcag	cctcctggat	acaagtgatc	ctcatgtctc	atccccctaa	10200
gtagctggga	ccacaggcat	gtaccaccac	gcacggctaà	ttttttattt	tttattttta	10260
tttttttgaga	cagtcttget	ttgtcgccca	gactggagtg	cagcagcgca	atcacgcctc	10320
actgcaacct	ccgcctcctg	ggttcaagtg	attctgtctg	ctcaacctcc	caagtgcctg	10380
ggatttacagg	cctgtgccac	catgcccgcc	taattttttt	gtatttttgg	taaagacggg	10440
gtttcaccat	gttgcccagg	ctgatctaga	acctctggcc	tcaagtgatc	cccctttctt	10500
ggcctcctaa	agtgtctagga	ttacaggcgt	gagcctctgc	acctggcctc	ggctaatttt	10560
ttattttttg	tagagacagg	ttctcactat	gttgccaggg	ctggtcttga	actcctgggc	10620
tcaagtgatc	ttcccacctc	agcctcccaa	agtgtctgaga	ttacagatgt	gagccactgt	10680
gcctggcctg	gaactcatta	ttgaagcatt	cactagatc	aactttgggg	ttacctggcc	10740
acatcctctg	acctacctat	aagggtatca	cagctaacgg	agcctctgtt	tctcagaatt	10800
taggcagaag	cagttcaatt	tatcacaaac	tactctatat	ccagcataag	tgcccaaata	10860
aaacaattgc	taaagttctt	taggcattta	ctgtttgtta	gttagatatt	tagtcctcac	10920
tacaaatctg	tgatacaggt	attattttta	tttaacccat	tttatagaag	agaaaacctga	10980
agctcagaga	tgctaagtaa	cttgtgcaag	gtcacacagc	tagtaaataa	agggcagagt	11040
aaagatttag	tttcacattg	gactccagaa	cctttctact	gggactcatg	ggaatagtgt	11100
ggatgtccct	gaccttcagt	ggcccagggc	tctcctgggg	gaatccagcc	atagacaaga	11160
caccagcgag	agcccaatcc	taagattttg	tttgtttgtt	tttgagacaa	ggtctcactc	11220
tgtcaccaga	ctggagtgca	gtggcatgat	caatgtctac	tgcaaccttg	atctcccagg	11280
ctcaagcaat	cctcccacct	cagcctcctg	agtagcttgg	actacaggtg	cacaccacca	11340
cacctgacta	atTTTaaaat	tttattttaat	taattactta	ctattatttt	ttgagacagg	11400
gtatcacttt	gtcacccaag	ctggactgca	atgggtgtgt	ctcagctcat	tgctcctcc	11460
acctcccagg	ttcaagtgat	cctcccacct	cagcctctgg	agttgcaggg	actgcagggt	11520
tgcgccacta	tgtcagcta	atgtttttat	tttttgtata	gatggggtct	cactatgttg	11580
ccagggtctag	tctcaaaact	ttggactcaa	gcgatcctcc	tgtcttggcc	tcccaaagtg	11640
ccgggattac	aggcataaac	caccacaccc	aacctctaag	gtgtttttgc	tgaatgtgac	11700
catgtcagag	gcaggaaagg	gaagcatcat	ggggtttagga	aaggaacact	gagcagggtg	11760
acaaagaaaa	tgggatcatt	ttgtgagtg	tgcgtgtgtg	tgtatgtgtg	acaattctca	11820
gagccagcct	ctcaggtggt	tgagaccaca	gtccccattt	cccagatgag	ataatggagc	11880
ctcagagagt	ttctgcagca	cagctagtgg	aattagaatt	tgaacccggc	tcttcagac	11940
tccaggtgct	tcacaaccat	ccgcaacctc	gtcatttgca	gtttaccttc	atgattttac	12000
catttccctt	tgccatagct	agtgttattt	acttaataat	tccttttgaa	tcagtctgct	12060
taaaaaaaa	tagcttcatt	ctaaagtgt	atattcttgg	aatatcggtt	ttgctgttac	12120
ccacccccac	acgttatata	tatacatgta	tgtttcta	acatatatat	gtacgtatat	12180
acgtgtatcg	ttttttgtta	ttttttttgt	tgttgttagt	ttttttttaga	tggagtctct	12240
ctctgtagcc	caggctggag	tgcagtggtg	tgatttccgc	tcactggaac	ctctgcctcc	12300
tgggttcaag	cgattctcct	gcctcagcct	ctggagttagc	tgggattaca	ggcaccacc	12360
actacaccgg	gctaattgtt	gtatttttag	tagagacagg	gtttcaccat	gttggccagg	12420

tgggtcttga	actcctgac	tcaagtgatc	cacctgcttt	ggcttcccaa	agtgtctggga	12480
ttataggtgc	gagctactgc	ggctggccaa	tgtatgtttt	taatacacat	tcaaataacg	12540
aataactatg	aaacctgaaa	aactgctcca	tgttacttcc	tgaacccatc	ttgagtgtctc	12600
acatgctgtg	cataccacat	attgggaaac	actgctttcc	ctggcttcca	agcccagctt	12660
aatcactgtc	ccatcctatg	cttcgcttta	tttgtctata	aatgtttggg	ttgggggttg	12720
atgccaaaga	ccttttctgt	tgtcattaac	atggacacag	ctctaagagg	tcttggcatc	12780
ttgggctggc	tctcctttta	gttcagaatt	tggattttta	tccaactact	cagagtgtac	12840
aagccttcct	tatgaatgaa	ctcgttggtc	aaactcataa	aaggctgtac	gataaaacag	12900
gaatgaatgc	atgaattgac	actaagtcac	tagcatttca	cgggaatgga	ttctccgtta	12960
gtggaagagc	acatgtcctt	tctggcactg	atgtgtgctt	gggaaactta	ctgagctaac	13020
tggcccatgt	aacacagagg	ccctttggtg	cagtggaaaa	ctgttgactt	tggagattat	13080
cttgagtttg	aatctgagcc	tgctgttaag	aagctggcta	actgaattgc	tttgcttctt	13140
ggacccttac	cattttataaa	atggggacca	ttgtactcac	ccttttaggg	tattgcatgg	13200
attaaatggg	attctctata	gaaaatattg	gcacaaagta	ggtgtaaatt	tgacgcctag	13260
tgggatttgt	tgtgagggaa	attgtcattt	gattatcaaa	gacttaggag	caggaacagt	13320
gtctaattca	gggactgcaa	atggaaatgc	cagctgaggg	caggcatttg	ctaataattg	13380
ggtaaagcag	ggcaggtgta	gaatagcaat	gtctgggaat	taaaagagag	gtgaggacgt	13440
gtatgacctt	gagaaggcaa	gccctggcaa	aaggggatgg	cctccactca	gctacagtca	13500
tgccctagatc	ttctaacttt	ttatttttat	ttttattttt	tgagacggag	tcttgctctg	13560
tcacccaggc	tggagtgcag	tggcgcgac	tccgctcact	gcaagctccg	cctcccgggt	13620
tcacgccatt	ctcctgcctc	agcctcccaa	gtagctggga	ctacggggcg	ccaccaccat	13680
gcccggctaa	tttttttttt	gtattttttg	tagagatggg	gtttcaccgt	gttagccagg	13740
atggtctcga	tctcctgact	ttgtgattta	ccctccttgg	cctcccaaag	tgctgggatt	13800
acaggcttga	gccaccgcac	ctggccgac	ttctaacttt	ttaaagagaa	gcaagacatc	13860
tggattttta	tgtgataact	cctgatttta	aactggcacc	caattataat	ttacaacact	13920
ataagggtca	acattgccag	cagagcaaaa	catgggtggg	ggcaactgct	ggtcacccgt	13980
gtgcagcctc	tgggtctaaaa	tcatctttgt	atttcttctt	gctttacgca	ttgtcccagc	14040
acagtgtctgt	tgtatagtaa	atatccagta	agtgggtgta	gaatgaataa	accaatgcag	14100
ataaacctgt	agagaggccg	ggcacagtgt	ctcatgtctg	taatctcagc	acnnnnnnnn	14160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	15960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16260

gaggacagggc	acagtggctc	atgcctgtaa	tcccagcatt	ttgggaggcc	gaagtgggtg	16320
gatcacgagg	tcaggagttc	aagaccagcc	tggccagcat	agtgaacccc	catctctact	16380
aaaaatacaa	aaaattagcc	agctacttgg	gaggctgagg	caggagaagc	gcttgaaccc	16440
gggaggtgga	gcttgcaagt	agccaaaaat	ttgccattgc	actccagcct	ggcgacaga	16500
gcaagactcc	gtctcaaaaa	gaaaaaaaaa	aaagatgaga	aagaggaagg	gagagaaaaa	16560
agagagagag	gaaagaaaga	gagaaggttt	tggagtcaaa	aagacttaga	aattccagtt	16620
cttccacttc	ccatggaacc	ttggcaagtt	gccttctctc	tttctctgaa	tctcacattt	16680
tgcctctgtg	aagtaggggt	ggtacctggt	ggagatgatg	cggagatgag	ggtgaggggt	16740
gtgttgacac	ctatgccccct	aggatgggtg	agagcttggg	agcactgaac	ctccccttcc	16800
cctcttgttt	cttcccccca	ttgtctccca	ccagctcccc	gggatctcca	cttcactctc	16860
tgggattcca	ccagcaggag	gctactcctg	gagttaaggc	gtgttggtca	gactggggca	16920
ttttaggggg	cataaataat	aattatgcct	ggacaatgga	cataacatct	agggccttct	16980
gaagcaaacc	agggtgtggg	gtacccaaac	aaggcagtag	gccccaggag	gcaggtccct	17040
gcagtcccag	cagagagcag	ggcacagggt	tgagaagact	gagcaaacct	cattatcagc	17100
tcctttgtcc	cccactctgt	cctggagcaa	tcattctggc	ctcttcccac	ttccccaaaa	17160
accaggtata	aaggctgctt	ctggccccctg	aagccagagg	cactgagagt	ggaggtctca	17220
gactcttgga	agggtagttc	ttttctggct	gccccaggcag	gaccagtgtg	ggcctctggga	17280
agaagcagca	cctcataggg	caaacacgta	ggaggcctgt	ccttaggaac	atcatagcta	17340
agcagacctg	tccccgcagg	ggcaggagtc	tgggctaagg	gtgatactgg	agagcagcaa	17400
cggagactgg	aagacaaatg	aaatttggtg	cctgagttat	ccctcccacc	attccttttc	17460
tagactctcc	agctcagggt	ctgttcatgg	caagaggaga	aagcaatctt	gtttgctctt	17520
taatcaaaaa	attaaacaaa	tattccctct	atactatgtg	ccagggggcta	tactagacac	17580
acaaagacag	ccccaaagaag	gacgggtggag	tagtgtctctc	gctaaaagac	agtagatatg	17640
caatgcctct	tgctcctgcc	ctttctcctg	ctgggaacag	ttctctgctct	tcactctgggt	17700
aagtctctcc	cttccctcct	catgcgtctt	tccctttttt	ccttttttct	acactcccct	17760
cccccgctt	ttatttgcac	tcatgaggcc	aggaccacag	ccttccctct	ttagctgata	17820
cagctcatct	ccggtgaagat	atcacttgga	ctcagaactg	taacctggaa	ctttctcttt	17880
tttgtttgat	ttttttttgt	tgttgttgtt	tttgtttttt	tttttgtttg	ttttttgttt	17940
tgttttgaga	cggagtctcg	ctctgttgcc	caggctggag	tgcaagtggcg	cgatctcggc	18000
tcaccacaaa	ctccgcctcc	cgggttcaag	caattcttct	gcctcagcct	cctgagtage	18060
tgggatacaa	ggcacatgcc	accacgcctg	gctaactctt	gtatttttag	tagagatggg	18120
gtttcaccat	atttgccagg	ctggtctcaa	actcctaacc	ttgtgattcg	ccgccccggg	18180
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgcac	ccggcaaact	gtaacctgaa	18240
ctttcagaag	gaaaaaccac	ccacctgtta	agatgaaggg	ctggtgactg	ccccaggctt	18300
ctcacacgtg	ctttctccca	ccttcaaaac	acacactcgt	ggtgtcggcc	agaagtcagg	18360
ttcttgtcca	tttgtgggtg	tgacccgaga	gatctctctc	tacctaacac	caaggaaatc	18420
ctccagtctt	gtcttcagggt	ggaattccta	ggaaagctcg	agcgacgttg	ctggagctgt	18480
ccacggtgct	ggaactagga	agctcttgac	ctgatggcag	gttacctctt	cttcccagag	18540
aatgatgccc	cccatctgga	gagcctagag	acacaggcag	acctaggcca	ggatctggat	18600
agttcaaagg	agcaggagag	agacttggct	ctgacggagg	aggtgattca	ggcagaggga	18660
gaggaggtca	aggcttctgc	ctgtcaagac	aactttgagg	atgaggaagc	catggagtcg	18720
gaccagctg	ccttagacaa	ggacttccag	tgccccaggg	aagaagacat	tggtgaagtg	18780
cagggaaagtc	caagggtcaa	gatctgcccgc	tacctattgg	tgccgactcc	taaaactttt	18840
gcagaagctc	aggtaaagtag	tagggaggct	actgcggagg	acctggggga	aaagagagta	18900
cattcagctc	tctgttccct	attcatttag	gctagtgttg	ctcaaagcct	cgcactcatc	18960
agaatcacct	ggagtgtgtg	ttaaaacaca	gctttctggg	cctcacctgc	acgacttctg	19020
atttaggagg	gctgaggtga	agcctgagaa	tttgcattta	caacaaatcc	ccaggtgatg	19080
atgatattgt	tggtctgggg	agaaccaccg	atttaaacia	aaggcttttg	tgttagaaac	19140
gcctgtgtta	aattctggtt	ctgcctttta	ttagctgtgt	tacctgggca	agttgctttg	19200
cctttcaaaag	cttttagcacc	ttcatttgta	aaacgaagat	atatagcacc	aacttcttag	19260
agttgtgggt	agcattaaat	gagataatac	atgaaaagtg	tttgggaatg	tcactgggct	19320
gtaataaaact	ctcaataaagc	ggtgggttata	attattatga	gtattatcat	ttcctgtagg	19380
attgtcctga	cagctaatta	agaagcaaaa	gataggatta	agggaggcaa	gtaggtttat	19440
ttttaacctg	aaaagggatg	ccgggctctt	gcctggagac	tcagaaactt	gaaataaatg	19500
agagggaatt	cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19620
gtagcacat	agccagaaca	tctagaaggg	gtggtaggag	tggggattag	aggttccagc	19680
tgagggaat	ggcacttgca	aaggcttctg	tgaaagtggc	taagtgtgga	ggttgagcat	19740
tcaggaaagg	agagcttcag	cttcagtgtg	gctggagtgc	tggtgtgtaa	gagaggtgaa	19800
gatgaggctt	ggaggctggg	cagattttgc	tccaaaagag	cttggtgaac	tgtgataagg	19860
agtttggtt	ttctcctact	aaggacaaca	gcaaactatt	gaagagttta	aatcgttcag	19920
tgacaatgac	acgtttgcgt	tttgggtggc	cactcgagct	gccagccagg	tagacagtgg	19980
cagaagatgg	aagataaagc	actaaagggt	gatgaggcag	gaagccagtg	aggagagaaa	20040
ggggacgatg	tgagtgcacg	taaatcattt	ggtgggttgc	tattgtgtgc	taagctctgt	20100

gctaaattct	tcacgtgtat	tatttcagct	aatccatcta	acaactctgt	aaggcaggta	20160
caatcggtcc	cagctgaaga	agctgaggct	ctcaaaagct	agtaacttgc	ctaagttcat	20220
gcagcatgca	agttgtccag	ccaggattct	aacttagaca	ccagaggcca	cttttaacca	20280
ctgctctagg	actgggggaa	atgggtcccta	gtgagatatg	tgtcgagttt	catatttcat	20340
tcaacaatat	tgttggcctg	ctacatgtga	agagctgtgg	aaagcgccca	aagttagtta	20400
gatccctatg	agcaagtggg	atgggggtgg	agtggacagt	aggagggctg	gaacacacat	20460
aaaagggtat	aagaaataac	aattaggccg	gccaggggtg	gtggctcacg	cttttaatcc	20520
cagcactttg	ggaggccgag	gaggggtgat	cacttgaggc	caggagtttg	agaccagccc	20580
ggccaacatg	gtaaaacccc	atctctacta	aaaatgcaaa	aattagctgg	gctgggtggg	20640
cacgcctgta	atcccagcta	cttggggaggc	tgaggcacga	gaatcacttg	aaccaggag	20700
gcagaggtta	cagtgaactg	agattgcacc	actctactcc	agcctgggag	acagagtgtg	20760
accctgtctc	aaaaaaagaa	aacaaaacaa	gtaggtactt	tctgccatag	ggaggattca	20820
taaactgcta	gtcctcaggt	gcattttttgc	ttatcagttt	taaaaatcag	agaatgtctc	20880
aaagaattag	gatgtcagct	tcttttgaaa	atgtgggcca	gaagcgggtg	ctcacgcctg	20940
taatcccagc	actttgggag	gctgaggtgg	gtagatcacc	cgaggtcagg	agttggagac	21000
cagcctgacc	aacatggcga	aaaccccgat	ctactaaaaa	tacaaaaatt	agctgggctg	21060
gtggtgcatg	ccttttagtgc	cagctactca	ggatgctgag	gcatgagaat	gcattgaacc	21120
cgggaggcag	gggttacagt	gaaatgagat	tgcaccactg	cactctagcc	tgggagacag	21180
agcaagacc	tgcctcgaaa	aaaagaaaaa	gaaaatttgg	aagatctgac	aacagttgac	21240
ctgcatttct	gctcggcaac	agcctgatgg	tggatgggca	gaggctcagt	tgtctgccaa	21300
acctcccatc	actgatgtct	tccctcgctg	tcacatctct	cttgacatgt	aggcatttgg	21360
tgtgtgcctt	ctgctctggg	tgcccagatg	aattggatgc	tatatgagaa	aacatttctg	21420
aatgtctctg	tggtaggcaa	cctcaaagat	cactggggcc	tccaatgatc	cctccttctc	21480
ggtattcatg	cctgtgtata	atcctctccc	ttgagtgtgt	actacacctg	gatactagct	21540
tctaataaac	agaacacagc	aagggtaatg	ggatgctact	tctaaggtta	aattacaaga	21600
gtgtaaagtc	tgtcttggtt	gtttccctct	cttgatcttc	ctctcattct	ctctctctcc	21660
ctctctctca	ctttcttact	gtcttgtcct	tcccttgggt	tactctgatg	aagcaagcta	21720
gcaagcatcc	atgttgtgag	ctgacctatg	aagaggccca	tgtggtggta	aggaactgag	21780
ggcagcctct	accagcaag	gaactgagtc	actcatcata	tgggtgagct	tggagacaaa	21840
tccttcccca	cttgagcttt	cagatgacgg	cagccctggc	tgatgctttg	caggcttgtg	21900
agagaccttg	agacagaaca	ctcagctaat	ctatacccta	tctcctgaga	tagagtataa	21960
tacatgtagt	tttaagctac	tatgttttgg	gataatttgt	tactcagcaa	tagataacca	22020
atacatatac	catgtacata	actgtttcag	ttgtctgaga	ctatatattg	tcattttaca	22080
cctacatcaa	gaatgtgtca	ggcaccattc	caggtaactt	gaatacatca	attaacagaa	22140
taggtaaaga	ggccaggcat	agggtcaca	tctataatcc	cagcactttg	ggaggccag	22200
gtgggaggac	tgcttgagcc	caggagttga	gaccagcctg	ggtaaaaatg	tgagacactg	22260
tctcaactaa	aaaaaaaaaa	aattagttgg	gcacagtggc	acatgcctgt	ggtgccagct	22320
gctcaggagg	ctgaggtggg	aagatcgctt	gagcccagga	gtttgaagct	ccagtgaacc	22380
acggtcacaa	aactgcactc	tagcctgagc	aacagaaaaa	gaccctgtct	caattaaaaa	22440
aaaaaaaaaa	aaaaggaaag	aaagaaaaaa	ataggtaaag	atccttgatt	cttgcctctc	22500
tggaaacttct	attctagagg	gggatgggtt	ttcacagtag	aagtctgtgt	tgacagcgct	22560
gtttaaagct	ccttcagcat	ctggggaaaa	ggttnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22740
tgaactcctg	ggctcaagca	atcctcctgc	ctcagcctcc	tgagtagctg	ggaatacag	22800
tgtgcaccac	catgcctggc	ttatttcata	tatatataat	tttatatata	tgtatatatta	22860
tatatataaa	tatatatata	atcttctgtat	ataaataaat	aaatatatat	atatatatat	22920
ttttagagat	agggtcttgc	tatgttgacc	accaggtcct	gaactcctgg	gctcaagtga	22980
tcctcctacc	tctgcctttc	aaagtgttgg	gattacaggc	gtgagccatg	gcacctaaat	23040
gagttatttt	taccacacga	agcataggac	atacatccaa	aaatgttctg	agctgagcaa	23100
gagcctggag	gcaagtgaat	ctgaactttc	ccgtctttga	agaaaccagt	ctctctccaa	23160
agtcacatag	ttagtgtcac	tcccccaag	aactgcatga	gctgggacaa	tcagagggca	23220
gtggaaggtc	tggggctcag	gggcgcccc	tgctgtctcc	ccagggtctg	ttcccttacg	23280
caagagcctc	tgctcccca	ctttcctgtg	gagcctcctc	accatgggca	tgaccagct	23340
gcggatcatc	ttctacatgg	ctgctgtgaa	caagatgctg	gagtaacctg	tgactggtgg	23400
ccaggagcat	ggtgaggcac	cgctgaggcc	cctgggggtt	gggggacacg	gcgggtcacc	23460
ctggctggagc	tccctcacc	atacgtttcc	ctacccacag	agacaaatga	acagcaacaa	23520
aaggtggcag	agacaggtag	ggctatgaaa	gcaggccctt	ggctcacgcc	cagcccactg	23580
caacccgctt	ctcagggggc	gggactcctc	taggcctggg	cccacccagg	taaccctttt	23640
gtgggatgta	agagtctggg	ttcagaggaa	ggctattttg	gtgctctctg	gcctccgctg	23700
gaaggggtga	tagtgtccac	tgagtgccag	ttcctgaccc	cactgccctt	cccacctctg	23760
ccagttgggt	tctactcctc	cgtcttcggg	gccatgcagc	tgttgtgect	tctcacctgc	23820
cccctcattg	gctacatcat	ggactggcgg	atcaaggact	gcgtggacgc	cccaactcag	23880
ggcactgtcc	tcggagatgc	caggtgacct	gcctgtacag	ggatggtgac	agcaagtggg	23940

caggcagtgc	ttttcatttt	ctctgtgcgt	ttacatccag	cagcttggtg	ctttctccca	24000
agaaccctag	gagatcaggg	gtacctcccc	attttacaga	tgaggaaact	gaggctagga	24060
agggacctgg	cttgcttaat	aataagaata	gctaatgcag	agtgcgtgact	gtgcacttgg	24120
caccttgcc	tgtttagtcc	tacaacacct	ctttgaggta	gatgcgttaa	tatcttcatt	24180
ttgcagttga	ggaaaccgag	gtacagggtt	gcacagttag	gtcattcacc	caagatcaca	24240
cagctttcag	tggcagcctc	cagaacctgt	gttataaggg	tacacgctaa	agtcttggtta	24300
gggctagaat	aggtagagtt	ggtatattag	atattttattg	ctgtataaca	aatcacccca	24360
aggcttggca	ttttaaaaca	acaaacactt	ctcatctcat	acagtttctg	acagtcagaa	24420
atcagggaga	gactcagccg	gctgattctg	agtcacagtc	tctcatgaag	acatagtcag	24480
gctgtcagcc	agggtcgag	tcatctgaag	ggctgactgg	ggttggagaa	tctatgtcag	24540
ttcaattacc	cccatggcct	ctccataggg	ctgctcagga	cacagcacct	gctttccctt	24600
gagcaagagg	gctaagcgac	agagaccccc	tatcttctct	cacataatct	cagacgtagc	24660
ataccatcac	ttctgttacg	ttctattata	ggcacagagc	aacctgata	tactgtggaa	24720
ggagactgga	caaagcaggg	gaataaccagg	aggcaggatc	cttgaggggt	gtcttggttg	24780
ctggagacca	ccattgaggg	tttttttttt	tttttttatt	gagacagtct	tgctctgtcg	24840
cccagggttg	agtgcagtg	cacgatctca	gctcactgca	acctctgcct	cccaggttca	24900
agcgattctc	ctgectcagc	ctcccagta	gctggacttc	acctggagt	cttgaaccca	24960
gattctgtga	ctgcttttgc	tctttttgtg	ttcatccaaa	cagtccctgt	ttatcctaag	25020
aggatgggag	aaagagactg	ggagagaagg	aatccagtg	gcctccctcc	ctgctagcag	25080
agcctggccc	tggcactgag	ccttcctcct	ctaccctctg	ctcctaattg	tgagggtccc	25140
ctagcagggc	ccttctgtcc	aggacacatg	ggcgcctgt	cctcacccca	gcctactgac	25200
ctctctcctg	ggctggcctc	agtgcccttg	attgtgcgg	agagaggaag	cgctggacag	25260
tcaggccaag	ctgctgtccc	caggagggca	tctgcttatg	tctagggcag	ggcacacctc	25320
ctgaggactg	ctgatgagag	acggtgtgag	agcttcccac	ttcccacctt	ctctccctac	25380
cttggttctc	aaaccttcaa	gtgtgcatga	gaatcactta	gtgggggata	tttgtccaaa	25440
tgcagatttg	cagatatccc	cgctgagatt	ctgagggccg	agatgaggcc	tgtgaatctg	25500
catgttaaga	aagcaccgc	tttgatgcgt	gtgtcattgg	gtaggggagc	aacactttga	25560
gaaacatgga	gctagagaac	gtgggtttct	atgggtttcc	catagaaaca	tggatttctg	25620
tgttttctgc	tgccctgaca	tcgaaggcac	atctgaagg	ggaggggcca	ggccaagaac	25680
cagggagtc	tgggaacgta	gaggcagcag	ccagtactcc	cccgactacc	tcaggagcgg	25740
ggttgtctac	aaatccatca	gaccacgcta	ctgcaagatc	caaaagctca	ccaatgccat	25800
cagtgccttc	acctgacca	acctgctgct	tgtgggtttt	ggcatcacct	gtctcatcaa	25860
caacttacac	ctccaggtag	ccaccttcat	ccttcccctc	tccctgcctc	ccgaggctcc	25920
tccaaaggga	tgggtccatcc	agcacctgcc	ttccaggaag	cgagttctg	gtcttctgat	25980
ctggatctat	tttccgggtt	ctccaggaag	tgtttctagt	agattgggtt	ggcgaggggg	26040
tgggaattga	ggcccagttg	gcctcttcgc	cctaccctcc	cttccctccag	cctccacaca	26100
ctctcctaac	ctcttcactc	tctctttttg	gttttagttt	gtgacctttg	tctgacacac	26160
cattgttoga	ggtttcttcc	actcagcctg	tgggagtctc	tatgctgcag	tgtgagtctg	26220
ttgggctgaa	atgccttctc	gagctttgca	accgtgatca	gagaacccca	gggaagggtt	26280
gggagggccc	caggcatccc	ctaattgcacc	tctctctgag	acctctgat	ggcagggagc	26340
tcacttcctt	aaaggcagcc	tatcctgctg	taattgactc	cccctgttgg	agtcttccct	26400
tagaggaagc	tgaataacct	ggcttgatga	cactttgggt	ctatgtctgc	tgtttgaaac	26460
tgcccccata	atggcctccc	ctccatgccc	acctgaaga	aatttcccaa	ggcgagccat	26520
ttgccttata	attttcctct	tcatgttggg	cagtcgccac	ttgcatctct	ctcctgggtt	26580
ccctgctgg	gcgctgctga	gggactctcc	cctgtgtatg	tgatggagta	acaggacatt	26640
acaataatga	tgacaaaatg	acaaccatta	tcaagtgtc	cgttgggtgca	ggcagcaggc	26700
aggatccttg	accatcactc	cctgagttca	gcctcactgc	agcgggtctg	gcagagggca	26760
gctctctttc	cttcatctgc	tcaagccaga	acctggagt	ttccttgatg	tttctctccc	26820
tcacactcca	tgttactctc	atcctcagta	cagccagcag	cagcttctac	acaccccaaa	26880
tctgaccttt	cttgtcacct	ccactgctgc	ctctccagtc	ctagccacca	acatctctag	26940
cctggattat	tgtggcagcc	tttagtctcc	cacatctgcc	ctggccccgc	tgtctcagtc	27000
tatttttaac	acaggggctg	cagtcacctg	tcaggacata	agtctcttca	catcactctg	27060
tgggtgtcctg	tctcatctgt	ctcagagtaa	aagccaaagg	ctttactatg	gcctaaaaag	27120
ccctgcaagc	tctggcccca	gcacttcact	cccctctagc	ttccccctct	ccattgttca	27180
ctctgccaca	gccacagtgc	ttcctagtgc	tccggaagtc	tcaagtgtgt	tccctgcttg	27240
gcactcttgc	atgtactagt	ccctgtttct	agaacattct	tctccagata	tctgcaaggt	27300
gcccaatctt	accttctctc	cttcttcagg	tctttccctg	actgtcctct	tctcagtgag	27360
gcctcccttg	gctgtcccat	gtacaattgc	aaacctccct	ctgcccgtct	ctctgcttgg	27420
tttttctcag	cgtttatcac	taacactctg	cctatctctt	gcttattgtc	tgaccgccac	27480
ctgctccatg	ggaatgccac	ctcctcgatg	gcaggaatct	gttgacttgc	ttgatcgtgg	27540
tatctccagc	acctagagca	gtgcctggca	catagtaggt	tctcagctaa	atgtttgttg	27600
acagaataca	gtggacagtc	ctgcgaggtc	aatgccatcc	ctgttattag	tgagggaagt	27660
ggggctcagg	gagtttgagc	cacttgccaa	tatcacacat	acaggaggtg	tgagaaccca	27720
gctcagtggc	cctgaagttg	gagcatttgc	cctcaaggct	ggggaccaaa	gagcccatgc	27780

aaagagcccg	aacgcttaag	caccaccctg	cctggccagc	ggggnnnnnn	nnnnnnnnnn	27840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27900
nnnnnnnnnn	nnnnnnnnnn	nnnnccact	gcgcctggcc	cattactttt	aatggcaaaa	27960
accacaatta	cttttgcacc	cacataaata	gttaccatgg	gctgagcatg	gtggctcagg	28020
cctgcaatcc	cagcactttg	ggaggctgag	ccaggcggat	cacttgaggc	caggagtcca	28080
agaccagcct	ggccaacatg	gtgaaacccc	gtctccacta	aaaaatacaa	aaattagctg	28140
ggtgtggtgg	cgcgtgcctg	taatcccagc	tattcaggag	gcagagggtg	cagttcactg	28200
aatcatgcc	actgcactcc	agcctgggcg	acagaatgag	actctgtctc	aaaaataaat	28260
aaataaataa	ataaatat	acatgtttt	gaccacctgt	tatgtgccaa	ctgtattact	28320
taaaaacaca	catgggaggc	tgggcacagt	ggctcacgcc	tgtaatcgga	cactttggaa	28380
gggcaagcgg	ggaggatccc	ttaaggccag	gagttcaaaa	ccagcctagg	taacacagta	28440
agccctgtct	ctacaaaaaa	taaaaaaatt	aactgggcat	ggtggtgtgt	gcctgtaacc	28500
ccagctcttc	gggaggcaga	gggagagggt	cgcttgagcc	cagcagtttt	aggttgcaag	28560
gagccaggac	caagacacta	cactccagcc	tgagtgcag	agcaagacac	tgccctctaa	28620
caaacaaaca	aacaaaagcg	acctgtgggt	aggtaggaac	aggctcatag	tacagatgag	28680
aaagcagagc	ttggagggtc	caagcgat	gccaagcaga	ggtccaagcc	gaggtctctc	28740
tgaatccaaa	gttaattccg	tctatcatat	caccacagcc	ctctctgccc	caggagaggt	28800
ctctgcccac	tccagccact	cacgtgtaat	tgacttcctc	aggggcagga	aaggcttcga	28860
tgggccagtt	gaggggtgcag	ttcagaaaga	taaggcaggc	caggccagac	caggtgaaca	28920
tgatgaccac	gaaggccaca	ccggcatcgt	agatcagctg	tgagaggagg	gggcaggccc	28980
gtgggggaga	ctgcctggcc	ccagacccca	ccaaggtaga	tcccaggcct	cagaggcctt	29040
aaagaagtcc	tcttctcccc	ttgtccttgt	gcccattttg	cagatgagga	aaccaagacc	29100
agaagttag	agtgcagactc	agaagaccca	tcattccttt	cttggagccc	ggctcaaaat	29160
cctagagagc	ttgaaatag	tctccacaaa	gcctgaagtt	gctggccact	ggctcaaaat	29220
atctctgaaa	tttccattat	cttaaaaaaa	tacatacatt	tttgccctatg	actccacaaa	29280
cattcatgtt	catgttcgca	caaaaatgtc	catttcatag	tacgtacaaa	ggaaaacttag	29340
tgctctaggt	ttaccggggc	taatcgtgtt	tatcctgccc	cttccctggca	cattccccag	29400
gggaaaaggc	aaacccagac	tgctcatgct	cagccttttc	tcacctttcc	caggtcctcc	29460
cacgtgcaac	aactgggggg	gttggggaga	gggaggtgca	agtgtctctg	ccaaggctcc	29520
tcaacccag	ggcaggtaag	ttctcaattg	aatgagattc	tgtgcaaatg	tgtagccct	29580
tcttatggaa	gaagctgatg	caccatctgt	cctcttgctc	tccccatacc	atctgaccag	29640
gataattaat	gtctgctctc	ccctcaggct	cctgtctcaa	cctttttctc	tgcatctctg	29700
gaccttggtg	ccttttctct	cctaggggca	ggacagagct	tcaaagggcc	acacccccaa	29760
atgtgtggag	gtaagatctg	gctcttcaaa	cactacttca	gttgaaaaga	agggagaact	29820
gcccaccctc	catgcctgcc	caccagaaca	actgatggcc	ccccaccca	tgctctctct	29880
caaactcctt	tggagacact	gagcaaaagt	accttcttta	gtactctttg	taaatgtcaa	29940
aacggtatgc	agtttggtac	tgcccaccgt	ggaggttgag	gagcatggca	tggctcaaag	30000
ggtcctttga	tatttgacag	aggaaattga	ggcccccatc	ttgcactgag	ctaaaacttt	30060
ggtcccttgg	cttcgaggta	caccaggttg	acctgtccag	gatccagcct	ggcataaact	30120
cactttgtga	ccttgagcca	aaccacccat	cctctctgga	aggtgtggaa	aatgtggcc	30180
ccaaaggctg	aataaagcca	gagagtcagg	gaccttgaac	gcatgtgaag	gggctggact	30240
tgattctgta	ggtgaagcta	aaccactgaa	ggtttttcag	cagtgtgtga	gccagttccc	30300
catctgagat	ctttctggaa	gtcacgtgag	tgacagagta	cagagaaaaa	gaatcagagg	30360
cagggagacc	agctgagaaa	gcttgctgtg	gcccaggaga	gagggggaag	gcctgcattg	30420
ggatgatgac	agagaaagga	gagcggagaa	gtcagacccg	tgggtcagca	ctagctgctg	30480
ctcactcggc	cccacccggt	tcttgtgtca	agacaaaaag	aaaacccagg	tggcctcata	30540
ccttgattcc	tgggaacgta	atggcagaag	aggcgtgaaga	gccaatcatg	agggccatta	30600
acgtggagcg	caggttccca	aacatgttgg	gcagctgagg	agggaaagca	gcacccatga	30660
ggtgggggaca	ccgtgaccct	tgcccagcat	tcccagccct	gctccataca	atagctccag	30720
gagacgcagc	agaaaagccc	caaggtaaaa	caaacagaaa	aatcaatgtg	ggaaaactgt	30780
ctctgcccc	tgccctacaca	gtcacagtgc	ccttttagctt	caaaaaggct	cccagacacc	30840
cctcagagag	acattttgtt	aattttgttt	aattccaggt	ttcccaagtt	tgttacgtaa	30900
cacctctgaa	aaacacatgg	aataggtgct	taagaaacac	tgatcttggc	tgggcgcagt	30960
ggctcatgcc	tgtaatccca	gcactttggg	aagccgaagc	tggtgggaag	cttgaggtca	31020
ggagttcaag	accagcctgg	acaacatggt	gaaaccccat	ctccacccaa	aatacaaaaa	31080
ttagcttaggc	atggtggcat	gcgcctgtaa	tcccacgtac	tccagaggct	gaggcaggag	31140
aatcgcttga	acctgggagg	tggaggttgc	agtgagccga	gatcgacca	ctgcacttta	31200
gcctgggtga	caagcggaga	ctatgtcccc	accccccaaa	aaaaaagaaa	agaaaagaaa	31260
gaaacagtga	tcttgtccaa	cccatttgag	atgagacaat	tgagacccag	ggaggaaaag	31320
tgtactcaag	ttcacagagc	acattaatgg	ctttctcccc	attgtcgttg	tcccagccct	31380
aacccaaggc	tgtgaccatg	gctgtgtccc	ggtaatatggc	agtgcctctt	aacctctctg	31440
ggtgacgtcc	cagcccagtt	tctgcctaag	caggacaaat	cacatcctgg	gaggtgaggg	31500
tggaaataag	ggagggaaact	gagccagggc	agacagtctc	cagaggaggt	ggctctgacg	31560
cagagcaggg	tcagaaccca	caccaggaga	gaatttaatt	gatcatgtgt	tccactcacc	31620

tgcctcagcc	aagccctcag	ggcaggggaa	ggcaaagtca	ggatgccctt	cgcacacacc	31680
ctcctctggc	cccaccatcc	tcccacagtc	actagatccc	acagctgaga	aggaccttag	31740
gatccgtaca	aagcctaaac	acactccaca	gagggggaaa	ctgagactct	gaagggaggc	31800
ctcaacagct	ctggtaaaaa	aggcgtttag	gccgggcgca	gtggctcaca	cctgtaatcc	31860
cagcactttg	ggaggccgag	gcgggtggat	tgcctgagct	caggagtctg	cgaccagcct	31920
gagcaacacg	gtgaaacccc	gtctccacta	aaatacgaaa	aaattagccg	ggcgtggagg	31980
cgtgcacctg	tagtcccagc	tactcgggag	gctgaggcag	gagaattgct	tgaacctagg	32040
aggcagaggt	tgcagtgagc	cgagatcgcg	ccactgcact	ccagcctggg	cgacactgcg	32100
agactccgct	tcaacaacaa	aaaaaaaaaa	atgggtgtta	aacacatata	actaaattat	32160
ccttccccct	tcccctgaag	tggctggctc	aggaaaaacc	tctaccact	caggcagagg	32220
ttttcctgca	ccctgcatcc	gtgaggcacc	actgccaaag	acgccaggga	aggctgccag	32280
gcctggagag	gggcaggggc	ccctccccct	caaggggcca	caaacgctgt	ctgcgcccag	32340
taccgtgggt	aaggcgaggc	cggccggcta	accccgggct	ggcggccttg	cagcgtgcgt	32400
ggcaacagca	gctgggcccc	caagactcag	cacgggacgt	cctcgtccaa	gtctgggccca	32460
agagcagcgg	cccagggggc	ggggccggcc	agagggagcg	gggagaggct	gagggggcgt	32520
gccagcgccg	gaccctgcc	ttggctggag	attacaggag	gcggggacat	agcagggagg	32580
agcgcctgga	caagccccac	ccggccggca	gggaggtct	gaggtcaaga	gccggagaga	32640
agggatttag	ggccctgggc	caagttgcac	agcagggaga	aggggctgcg	cagaggggcg	32700
gggagaaagg	gatccgcttc	cttccttttag	agctgtgaaa	tgtccccggt	tggaaattaaa	32760
ggcggctgct	ggggagaggt	gaaattcagc	caaaaccacc	cagtcaggca	gcccttctca	32820
gagataaaca	gtccgagcca	gcccggccag	gaaccttccc	ctccaacctc	cctaagcctt	32880
taacactcct	aagcctttta	cgcgttttag	cactcacata	aataaacaca	ctttgagcaa	32940
cacacataca	ccactcacca	catgtaatag	gtcaagccat	gtgcacgacg	agggtgtcgac	33000
aatttcatat	ggttcaacct	agtacactca	caaacacacc	taccaactca	tggctttcac	33060
agggacgggg	tcacacaccc	actctcccac	gacatggcaa	gcgtgcacac	gctatctcaa	33120
gctgtccct	ccccctcaag	atcatgttac	ccagttttat	tttcttccca	gcacctatga	33180
cgactgacat	aattttattag	tttacttggt	tattgggtta	tctgtgcccc	tcacccccaa	33240
aatgtaacct	ccagcaggga	ggatgactcg	gtcagtcctg	attgtgctgt	agtccaggac	33300
ctagaacaga	gctccatgga	cattcatggg	ctctgtacac	acaaacacac	acattaacat	33360
acaccgcagc	acacagcctc	atccacacac	acacagcctc	acacctgtct	tttgtagcca	33420
cctgcacagt	ttctcacaca	ctcacttgat	ctagtgtatc	gcgtccacag	gccccctccc	33480
cagcccactc	atactgccct	caccccactc	actctgccct	caccccactc	gggggaaactc	33540
tgctgccagg	ccaggcctgt	gacactcacc	gtgagtgaag	tgaacgttag	gcagatgccca	33600
ccaaagccat	tcagggacag	cgccagggaat	atcaacggag	acagagctgg	aaaggggaaa	33660
gcagcagatg	agggcatttg	gggagctgtg	ggaagccaag	ggcgggagct	gggggtaaaaa	33720
tccgccttca	tcccacctat	tcttttcttg	tggggccaca	agaggacaga	caactcacct	33780
tccacgtccc	gggagggcag	ggccatgagg	gtgcaggacg	cagtgaagca	ggcactgtgg	33840
agacacaggg	aagggcgagg	ggttggcctg	tgagcacccc	ccctccccct	ccccctgcagc	33900
acggctccctg	tcctcccgtt	ccccatagcc	cagccacctc	acctgccaac	cagccgcacg	33960
ggtcgggggc	caaagcggtc	catgaggatc	cccagtgcca	gggtggtggc	gctgagcacg	34020
aaggaaccaa	tgggtgaagcc	caggttgagc	atctcgtcct	gctggtcaca	gcctggccac	34080
ctgcgctgct	catcctgggt	ggtgttggtg	ctgctctcag	ctgaggaggg	ggaagggagg	34140
gctcagcaca	tgacaccagg	aacagctggg	cacaggagac	agcagcccac	agtcaggcgg	34200
cctgctttca	aatcccattg	caagtgcctt	tgggggtacc	ctagagtcac	atctcctctg	34260
atggggctgc	tccagaaatg	gcagccatta	gtacctgacc	ctgggagagt	cttgtgcaca	34320
cacagcctga	ggcttcaact	agctcaaagt	aaatactgga	cataaaaagta	tttactaagt	34380
tgtaatatgc	actcagtgct	caagcttagg	gggttggtga	cccccaacaa	gaagtgtccc	34440
cataatctaga	ggcaaaggca	aaggcagtg	gtggtactct	aatggctata	acaagaattc	34500
attaaaatgg	ccggcgctgg	tagttcatgc	ctgtaatccc	accactgtgg	gaggctgaga	34560
caggcagatc	gcttaagcct	acaagttaga	gaccagcctg	ggcaacatgg	taaaacccca	34620
tctctaaata	aaaaaaagaa	atthagaaag	aacactaaaa	cttagaggaa	gctttcccga	34680
taaatgatag	tctgataaaa	taatagctaa	tacttattga	gcacttaact	atgctccagg	34740
cactgttata	agtcagttaa	taaagtatcc	cgttccctag	gtgatgaagc	tgaggcacag	34800
aatgagaacc	aggcactgcc	ctccagtcct	ctctagaagt	ccacttgagg	gacttgtcct	34860
taacggtaaa	ctgccaaact	ggagttgtga	caagttaagg	agaaaagcta	gtgataggag	34920
acaaagggct	gcttcgcttt	actcaatgct	cannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35460

gaaaggaaaa	tgcaaagggg	gaaggggtca	cacttgggga	aggtttcaga	caataccgag	35520
tggaaagggg	gatgccaggt	gtggggagta	acagatagag	gaggcaaagt	gagtggagac	35580
caagccagac	cggggaggag	ggggccacag	ccaaggtgag	acaggtcagc	agccagaaac	35640
cgaagcagac	acttgcaggg	tgcaacccgc	cctctcttcg	tggcaatctg	agaccgagga	35700
cgtggagagc	ctggagagcc	cccaaccttg	tttctggggg	gtgggtcaga	gaggaagcct	35760
ctcatccccg	ggcaccagcg	gccttccccg	gaggctcaac	acgcagatac	ctggataggc	35820
gtccatcact	cccccgcca	gagccacca	acgtcctcgc	aggtccgacc	ttgtccctcc	35880
ttctacccca	cagtccccag	tcttagctca	tctgcataaa	gctccaatta	acatgttttt	35940
ccttttgctat	ttgcgatccc	agaactcggt	ccccaccccg	agcccgtttc	ccgccgcttc	36000
ctcgcccctg	ggagggcgjc	cccattaacc	ctcgcgacce	gggccgctcc	tggcggtcct	36060
gaccccgcca	ccccgtccc	cggcggggtg	ctgggggtga	ggggcgcgcc	ggggcgcgaga	36120
ggattgcgcg	gcagggtctg	ccacagggca	gaggccaggg	ctctccggga	aaaaggcagg	36180
cgcataatag	cccccttttc	tgggaaaaga	cggggagggg	ggcttctcct	gggagactcc	36240
aggcttcgaa	attcctcggt	ccctatcctc	cggcccccg	acccctcctc	ctccccgcca	36300
cgcacctctc	cccctcccc	gccatctggt	ccactccgca	gcgccgagac	aaacacgggt	36360
ccagctcgct	tccgcccctg	cccagcccc	tccccaaagg	ccggggagtg	ggggagtgag	36420
cagacgccct	tctcctagga	ggcgggaatt	tctgcctcca	tctcccaccg	gggtccgggt	36480
ggccagagct	aagcttcgag	acccccacc	aaccaccacc	accgttgcca	gggccggtga	36540
ggctgcagat	aacgcttgca	aggacgggag	tcggggaggg	tgtaggcgga	gtttaaagga	36600
cgggcagagc	aagccccggg	aagaggcagg	ggttttccct	cccgggtcgc	cgccccccgc	36660
accctcggag	ccagccgcag	ccacgcagcg	ccgcctgccg	ggcacaccaa	ggacctggcg	36720
cgcacgtggc	gcttaccccc	acccccgggt	ccgctcctgg	ctcgcgctca	gcctccccag	36780
actattcgca	aattgaggat	ccgggacaca	gagtgcagag	accccgga	gcctactgaa	36840
agccagccga	tcccgtgggt	gggtgctagc	caattctgat	tttgtacttt	acaaaacaa	36900
aaaaagtcag	tgttggaaagt	cgggagtctg	ggctcagagc	agcagggatc	tgcgatgtga	36960
ctttgccaa	tctccagacc	cctgaggaca	ggttttccta	tctgaaaacg	gaggggacag	37020
tctctcttat	taacttctca	agagaaacaa	agacaaaggg	agggaaaatg	gcttagctgg	37080
aatgctgtct	tacagagcca	acctttggag	gtgggggaga	tggccaaggc	ctctgaggtc	37140
actcttgccc	ccaggagcag	ctgagaaccg	gaaaagaagt	tgggacctcc	tttctgcaga	37200
gctatccttt	ccacagactg	ccgagggtcc	aaattgagct	ccaccaccta	acactgtgtg	37260
cccttggttg	tgtgccttaa	cctctctggg	cttgtttcct	acagcgacaa	gaaagaatga	37320
caacaccaac	ctcttaggct	atagtttggg	taaaatgaga	tagctgtgta	gaacagacag	37380
atcctaaacc	aattgttagtt	ttcccttcat	ttggggactt	gctctaacct	ccagggttta	37440
tgtcccagag	gcacaagcag	gtgcagggtc	ggataaataa	ggtatgtctt	tctgcaggat	37500
ctcttgctct	cactgatggt	gtcttctctt	gatatagata	atttttaaagc	ttcacgttat	37560
ttattttat	acttttaaagc	ctcactttta	tgttaaaggt	aaatgtaaat	atagtataac	37620
aaggaagctc	aaaattttgca	taaagtttta	agataaaata	ggagactcca	aaaaagtgtt	37680
actttcggca	ggccctaggg	atgctatggt	gggaagtgtg	agtcatacct	tagcattctt	37740
tctaaagcat	tctgtcctaa	tctctgttat	gggaaaaagc	cagcttctctg	gatgtacccc	37800
aaatcctggg	aagttagggg	caggagctgg	actccctcca	agcactaagg	gcagggcag	37860
ggtgggaaca	gggaggtgag	ccagacagcc	agaggcgaa	gggctggcat	gccaagcgtc	37920
ctagttaatg	cccagctgag	cctgggtgaa	gaaggatggg	ggtgtgggga	agacaccccc	37980
caccaaccgc	caaagacagg	cgcacaccag	ccagtctctc	acttcccttt	ttatttccct	38040
taagacttgc	aagcagcagc	accagagagg	gaacctgccc	tccctggccc	ggaaggggcc	38100
gacccccaac	ccctaaccac	ggacacagct	ggcacctcag	gcccccttcc	ttctgaaagg	38160
agggctgtgt	ctctctcaca	ttcacacata	cacagacaca	tgcattgtgtg	cacactcatg	38220
gcacatggga	cctcaggggt	agcctgtttg	ccgatcccc	caagaggtag	caggaggcag	38280
accgctagaa	ggagataaga	ggcacccctg	tctcctccaa	cccaaggagg	aagaaagctc	38340
aacccctcta	ggatagggac	tgtcttcagt	caatggagcg	ttgacttagg	gggcgttttt	38400
gaaggttttt	tttctctcct	tttgagctct	ttacaaaaat	agaacttctc	ttgggtattta	38460
taaatctacg	ccatggctc	tatgtgcatg	ttacaggtag	aaaagccata	tggggcactc	38520
cttttggttg	ctcaggcctt	gattgcctgt	catccaggte	ccttggtctg	agaagtctat	38580
gcggtcacct	cagagccgct	aagcaccttc	agtgggcccc	tcccattggc	ggcgactctc	38640
tgctggagcc	gggcacggta	atagaagagg	taggaaggca	acaggaatcc	caggagttag	38700
aatagcagga	ggcccagatt	caccttttag	gcaaggagag	agaaacagag	tcaagttagt	38760
agtcactctg	ccttagcctc	ccacagggag	gagaaggcgg	ccatttttct	ccaggtcctg	38820
agccagaata	aatacagcta	gtacttatta	tggtgagtca	ttgttcacac	agtatctcac	38880
ttaatgttca	gcaattctgc	aaagtggctg	agtgagactc	tctcagggtat	aaacaagtggc	38940
agggcctggt	gggtgcccac	accatatggc	actcactagg	taggtatgag	gaaggcacag	39000
cactgtagga	gtctgggctg	gtcaggctgc	tcccgaatg	gggccttctg	ggctcacccc	39060
tctgaccttt	ggagatgtta	accaatggga	tcccgttcag	ggtggcgaga	ggaggctctc	39120
agacacagtt	caagggaactg	ggatgcacag	cctggtggac	agaaggcttg	gaaggcccag	39180
gacacgcggg	ctctgactcg	gttcacatcc	cactctgcat	tactcactgt	gtgactttgg	39240
gcaaataatg	gcaattctta	ctgagtgcct	ccttctcagg	gctgttgtgg	cgaagatgta	39300

agttaaaaaa	aagtatgcat	catgcttagc	acatagtgag	tgcttggtaa	atagaagcag	39360
ttatttcata	acaattcttt	gggaggagg	tttacgtgtg	ggtggcccca	cagggcagat	39420
gaaagatcag	cgtcaggag	gcagatgagt	tcaatgtaag	gaaaagactt	actaacagca	39480
gcagggtctg	ctcgtgcagg	agtgggtgcc	ctaccactga	gggtatctaa	gctaagaggg	39540
aagggtcccc	tttcagggggt	gctggagaca	ggatcccaca	ctaggtagaa	ctggattgga	39600
ccaatggtgc	ctgaacacag	gcccaagagt	caggactggc	cacttcacaa	agcacctgga	39660
gtttactaaa	aacagactcc	taggaggtca	ggcactgtgg	ctcacgcctg	taaccccagc	39720
actctgggag	gccaaggtga	gaagatcatt	tgaggccagg	agtttaagac	tagcctgtgc	39780
aacatggcaa	gaccctgttt	atctgtacaa	aatttttttt	taaaaaatta	gccaggtatg	39840
gtagccatca	cctgtggttg	cagctactca	gaaggctggg	gccggaggat	cgcttgagcc	39900
caggaatcag	aggctgcagt	gagctgtgat	tttaccaccg	cactccagac	tgggcaacag	39960
aacaagacac	cttctctaca	aaaaaaaaaa	aacaataggg	ccgggcgcgg	tggctaaggc	40020
atgtaatccc	agcacttttg	gaggctgagg	agggcagatc	acgaggtcgg	gagatcgagg	40080
ccatcctggc	tagcacggtg	aaaccccgtc	tctactaaaa	atccaaaaaa	aaaaaaaaaa	40140
ttagctgggg	gtggtggtgg	gcgcctgtgg	tcccagctac	ttgagaggct	gaggcaggag	40200
aatggctga	acccgggagg	cggagcttgc	agtgagccga	gatcgcacca	ctgcactcca	40260
gcctgggcaa	cagaatgaga	ctccgtctca	aaaaataaaa	ataaaaaata	ataaataaat	40320
aaaataacaa	taaattaaaa	acaaaaacag	actcctacgg	tcaggctgag	atatcctgat	40380
tcaggggact	ggggaatctg	tatttttaac	actccgtgag	gggttctaaa	aggcagacaa	40440
cttggaiaacc	tgcagattag	agacctctga	ggtgcctctg	gctgagatga	gtgagggatg	40500
gcaccacata	caaggcccta	cccctgcccc	caggagagtg	gctcctgctc	ccccacaccc	40560
aaccctcgct	ctcaccacga	agggtctctc	tttcaggggt	cccaccatcc	ccatgaaaaa	40620
tggctgtgta	agcaaggcga	acacagcact	ggtgagggac	tgcaaggcctg	tcagctgtcc	40680
aaaaggggtt	ggatggggaac	ctgtccccaa	aacgggagat	caaaggggtg	tgggggcctt	40740
tcagcccagg	caagaacttt	ttcttttctc	tcccaacatg	ggnnnnnnnn	nnnnnnnnnn	40800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40920
aagaaaaaaa	aatctttaaag	aataaggata	taaagaaaga	aaatatTTTT	gtgtagctgt	40980
tcaatgtttg	tatttcaagc	caagtgttat	tacaaaaacag	tcaaaaagttt	ttaaaaattt	41040
aaaagtttat	aaagtaaaaa	agctaagtaa	gctagggtta	attttttttat	cgaaacaaaga	41100
aaaatatctt	tgtataaaact	tagtgtagtc	taagtgtaca	ttgttttttat	tttattttatt	41160
ttttattttt	ttgaaatgga	gtttcactct	tgttgcccag	gctggagtgc	aatggcatga	41220
tcttggtctc	cggcaagctc	tgtctcctgg	gttcaagcga	ttctcctgcc	tcagcctccc	41280
aagtagctgg	gattataggc	acccgccacc	atgcatggct	agtttctttg	catttttttt	41340
ttgaaatgga	attttgctct	ttgaccacag	ctggagtgc	atggtgcaat	ctgggctaaa	41400
tgaacctctc	acctcccagg	ttcaagagat	tctcctgcct	cagcctcctg	agtagctggg	41460
attacaggca	tgcaccacca	cactcggcta	atttttgtat	ttttagtaga	gacagggttc	41520
tcaactaaag	agaacctatg	tggccaggct	ggtctagaat	tcctgacctc	aggatgatcca	41580
cccacctcgg	cctcccaaag	tgctgggatt	gcaggcatga	gccaccatgc	ccagccagta	41640
tacagtgttt	ataaagcctc	cagtagtgta	cagcaatgtc	ctagaccttc	acattcactt	41700
actactcat	cactcactca	cccagagcaa	ctgccagctc	tgaagctgc	atgcatgata	41760
agtgccttat	ataggtgaac	cattttttta	tatttttatac	tatattttta	ctgcaccttt	41820
tctatgatta	gtacacacaa	tgcttaccat	tgtgttacaa	ctgcctacag	taatacgtac	41880
agtactatgt	atgggtttgt	agcctaggct	ataccatgtt	gcctacgtgt	gtagtctgt	41940
atactgtcta	gtttgtacac	tctatcatgt	ttgcataaag	ataaaatcac	ctaatagcac	42000
atttctctga	gtgtattcct	gttgtaagc	aacacatgta	taaacattta	caagaaatag	42060
ctcaaatTTT	tttttctttt	gatacagggt	cttgctttgt	caccacaggct	ggagtgcagt	42120
ggcgcaatct	cggcgcaactg	cgacatctac	ctccccggtt	caatcgattc	tccggcctta	42180
gcctcctgag	tagttaggac	tacaggcagc	caccacacag	cctggctaatt	ttttttgtat	42240
ttttattaag	agatggggtt	ttgccatgtt	gctcgaact	gtctcgaact	cctgacctca	42300
ggtgatctgc	ccgccttggc	ctcccaacat	gctgggatta	caggcatgag	ccaccatgcc	42360
cagccattac	gttttttttg	ttgtttaatt	tttttttttt	taagagacag	attctcactc	42420
tgtcatcaag	gctggagtgc	aatggcacia	ccatagctca	ctgcagcctc	caactcctgg	42480
gctcaaggga	ccctcctgcc	tcagccttcc	cagtaactga	gactacaggt	gtgagccacc	42540
atgctcagct	aattattttt	tatcttttat	tttttttaga	gggggggtct	ttctagcttg	42600
ctcaggtttg	tctcaaaactc	ctgggctcaa	tcaattctcc	tgctttggcc	tcccaagggg	42660
ctgggattac	aggtgtgagc	ctgaaaacct	tctagtgtgg	aagtggaaaga	taggccagg	42720
ccacttatgt	tttcaagtta	agcaaggttt	aggtcactta	tgaagcctga	ctagttttgt	42780
ttgcttaagg	gatctgcagg	cctgacctcg	gttttcattt	gttttaacag	tgtctatgtg	42840
tatgtgtgtg	tttatgtacg	tgcattgatg	ggggaaagct	cagaaatcaa	gtaagccaaa	42900
cacaaacatg	taattataag	cagggataaa	ttctatgatg	aagaagtatg	ggccacggga	42960
gagtcactgt	gccagtcctg	tgatcaggaa	caatgtcctt	tgggaagtga	catttgagcc	43020
atgccttgaa	gtacggtagg	agttggttag	gggtgagcca	gtaagaccca	gagctggggc	43080
ttcctgcaca	agctcagctg	ggcactgagg	accagtgga	ctctgctaca	gggcagtgag	43140

gagcagaaag	gctgaggaag	gctgggtgtg	gtggctcaca	cttghtaatcc	cagggctttg	43200
agaggctgat	gggggaaat	cggtagagct	caggagtgtt	agaccagcct	gagcaacata	43260
gcaagactcc	atccctgtaa	aaagctttta	aaaattagct	gggtgtggtg	gtatgcatct	43320
gcagtctcag	ctactcaaga	ggctggggta	aggattgctt	gagcctagga	ggtggacgct	43380
gcagtgcgcc	acgattgtgc	caactgtactc	caacctagga	gacaaagcga	gatcctgtct	43440
caaaactgaa	tgaataggct	gtgtgcggtg	gctcactcct	gtaatcccag	cactttttgga	43500
ggctgaggtg	ggtggatcac	ctgtgattgg	gagtttgaga	ccagcctggc	caatatggtg	43560
aaaccgcata	caaaaattaa	ctgggcatgg	tggctcacat	ctgtaattcc	agctactcgg	43620
gaggctgagg	catgagaatg	tcttgaaccc	ggggggcaga	gggtgcagtg	agctgagatc	43680
gcaccactgc	actccagcct	gggagacagc	gagactccat	ctcaaaaaaa	aaataataat	43740
aataacaatt	aaaaaaaaat	taaaaggcca	gggagcactg	gcagcctgtc	caagggtttca	43800
ggtcacttta	gtaaaggagg	aacaatggct	cctcccagga	cctctgggat	ctcagcattg	43860
atagcacagt	catggaaatg	ctagggccca	ggcagaccat	ctcagggaaa	acaagtggct	43920
ctgccctgcc	ttggccactt	cctggccctc	tgcattgccc	agggtctcag	caccaagctg	43980
ttctcagtga	gtagctctca	tttagtgcca	gggtctctcg	gcttacatcc	tacgatgacg	44040
atggaatgca	taaaagatgg	ggctgtgata	gcccagagct	aggggtttga	atctcatgag	44100
atgttcatgg	agccctggga	gggagctcag	tgcaagtcca	tttctctttt	ttggttgaga	44160
tggggctcag	aggaggaagg	acttggtcaa	agacacacag	ggagtgtttc	agtgtgggac	44220
ggaggtttat	ggagaaaagg	tgaccatcca	aggcttgga	aaagatcatg	acttcgacca	44280
gcaagcctca	actctgtaga	cttgggtggg	gccaggccct	cccaaacaca	cctgacaggt	44340
gtctgtgggc	ttggggacat	tgctgcctcc	cttctgtctg	atgctctgct	gtccctctcc	44400
catgaagcgt	atctcttcgc	cgcccccat	ccttctgtag	agaggatggg	ttctctctctg	44460
accaataatg	aagatcttta	gtaaagttct	cttttttttc	attttctgaa	agtcctctc	44520
ttgagaactc	aggacaagtg	agtcagggcc	aggacaaaaa	acagtgtggg	acgagtgtgg	44580
tggctcacgc	ctgtaatccc	agcacttttg	gaggccaagg	tggcggatca	cttgaggtca	44640
tgagtttgag	actagcctgg	ccaacatggt	gaaacctcgt	ctctacaaaa	tacaaaaatt	44700
agccaggcgt	ggtggtgcat	gcctgtaatc	ccagctattc	gggaggctga	ggcaggagaa	44760
tcacatgaac	ccaggaggcg	gaggttgca	cgagctgaaa	ttgggccact	gcactctggc	44820
ctcttgga	cagagccaga	ctacctctca	aaacaaaaac	aaaaacaaac	gacaaacagt	44880
gtagactttg	tgtttttctc	aaaagcactg	tcaagccagt	gcccgcagca	gtgggcctag	44940
acacctccag	tcttgectca	gggtcagttt	ccagcctccc	tggacacttc	ccccaggtat	45000
gtgtactttt	tgattgtcct	aaatccagag	tctgtggcct	gacctgggtt	gtcacagctc	45060
tcagtccctc	cccattcccga	atcccaggga	gccgcagggt	tgtgcagaag	aggcacacca	45120
cactcaatac	atcttgcatc	ctcgtggtg	ccaatccatt	ggcttggtga	tgtacagact	45180
gagcctcatt	atagccgttc	gttctgtgtg	acctttccag	atcaatctgc	cagcttggct	45240
tctccgagtt	tcgcttgtca	gcattttctc	aatcccatca	tgtacttttg	acctctttgt	45300
tgggtggcct	gcttttatctg	aaattttccag	atttgacttc	aggtctctcc	tttgtccctt	45360
aatatggcct	aatggtggac	cctgtcaggg	gtagagaaaa	tattgaggag	ccctgacttt	45420
gaggtgcaca	agtttagagg	ttagacaagt	ccagccacaa	ccagcccaag	ctgcagtgtg	45480
gggaggcctg	tccagctgct	ccacgggtga	gggtggagca	tacaggaagg	cttccttctt	45540
gctgcagccc	aggtgttctg	gctgccctag	ctgcctggct	ttggtagaag	aaagaaaggc	45600
tctgtctctg	acttgtcaac	taatggcact	atgagattgc	acataattaa	cctgggtctg	45660
ctcttccaaa	agccttgggc	ctctgactgc	aacatggagt	ctgggtatca	ctccccatcc	45720
ctgcgcctc	agcctgctct	ggcgctaggc	gtgtgcctaa	tcacttaatt	tctctgtgct	45780
gcctcttagg	tatcacttcc	cctgatccca	aatacttacc	aggtgtggga	tgacacctga	45840
ctagttactc	cttggaggta	tctgcttctc	accggggact	ccgaaaccaa	acgaaaagca	45900
aggccaagcc	cagcctaaag	gacgcttcct	acatgacttc	aggcttgcgg	gggctggagc	45960
gtgggggtgg	caatggagtt	ggggggggct	cagggagggg	atgtggaagt	gctttgtctt	46020
gcaaaactta	gagaaccgtg	taaataggag	tgattattct	gtcccttccc	ttcttttcca	46080
acaggaatca	gcattcccaca	gcccattgta	aatggaaact	aatggaaact	gaggctccgg	46140
gaggggtata	gggaggagcc	agcagggtct	tgagttcata	ttagtgccct	ttcctccata	46200
ggcacatctg	tgttttcttt	tatttttatt	tgaatttaaa	tttttttttt	tttggcagag	46260
tcttgcctctg	tcgcccaggc	tggagtgcag	tggcgcggtc	tcagttcact	gcaatctccg	46320
cctcctgggt	tcaagtgatt	ctcctgcctc	agcctcccga	gtagctggga	ttacaggtgt	46380
acaccaccac	acccagctga	tttttgcaat	tttagtagag	acagggtttc	acaggtgtgg	46440
ccaggcttgt	cctgaaatcc	tgacctcaag	tgatctgcta	gcctcgccct	cccaaagtgc	46500
tggtattata	ggtgtgagcc	actgcgctcg	gccacatctg	tgtttttaaa	gagaggaaag	46560
gggataatgt	gcatttttgt	gaagcttggg	ccgttttgtg	ctaggactct	tatgatcttc	46620
ataagttttc	ccccaggagg	gacactgttc	cacttaggga	gtcaggaccc	ccagtcctta	46680
caagattcag	cctctcaaaa	tggagacagc	agttccaggc	ctgggctggg	ttctgttcac	46740
actaggagag	ggcaagttag	tgggtgtttg	gatgtgggga	agtattatga	aaacagagat	46800
gtcceaattc	ctagtgatag	gaaaccatta	agctacttgg	catcttaaaa	ccaagagcgg	46860
ttcaagttct	gagattgtta	acacacctta	caacaccgcc	gccgttatta	ggaagaagct	46920
ctgtttgatg	acgtccca	ctgtgggtac	ctttatgaac	aggaatttgc	tttttcaaat	46980

cccagagaag	taagattaaa	gttggtgtgt	ctccatcctt	gaaaaatttg	gttttaggg	47040
gaattcaaga	atgactgacc	atacagaatg	gggagcaaac	ttgggaagaa	agaaggcaca	47100
gttcagagct	ctcccaatag	tcacccctga	actgcacccg	gaccatcagt	tatctctgtg	47160
ggtagagctc	aggaatctaa	aatccatttt	aaaattaaag	tatatcgggg	ctgggcgagg	47220
tggctcatgc	ctgtaatccc	agcacttttg	gaggccgagg	tgggaggatc	acgaggctag	47280
gagtttgaga	ccagcctggc	cacatgggtg	aaccccgctc	ctactaacia	tacaaaaatt	47340
agccaggcat	ggtggcagac	acctgtagtc	ccagctattc	ggaaggctga	gtcagaagaa	47400
ttgcttgaac	ctgggaggca	gaggttgtag	taagccaaga	ttgtgccact	gcactccagc	47460
ctgggcaaca	gaggagact	ctgtctcaaa	aaaaaaaaaa	aaaaaattaa	agtatgtcat	47520
acatactgtt	acaggcacag	accttaagt	tacagcccaa	tgaaatttta	cacatctata	47580
cagctatata	actaccacct	atatcaagac	acattccagg	aactcagact	ccatcatacc	47640
cctcctcagc	agaggtaaca	gacccacacc	tctcctgtc	cggtggtaat	taaccactat	47700
tctaactttt	ctatcaatta	gttttgccca	ttcttgagct	tcacacagat	atacattgtc	47760
aggcatgatg	actcatgcct	gtaatctcag	cactttggga	ggccgagacg	ggagtatcac	47820
ttgagccag	gagttggaga	ctactctgga	caacatagtg	agacccccga	ctctacaaaa	47880
aaaataaatt	agctggtcat	ggtggtgcgt	gcctgtagtc	ttagctatct	gagacgtgga	47940
gagaggagaa	tctcttgagc	ctgggaggtt	gaggctgaag	tgagccgtga	ttgcaccact	48000
gcactgcagc	ctaggtgaca	gagtgagatt	ctgcctcaaa	aaagaaaaaa	tatggccggg	48060
cgcggtggct	caagcctgta	atcccagcac	tttgggaggc	caaggcgggc	ggatcacgag	48120
gtcaggagat	ggagaccatc	ctggctaaca	cggtgaaacc	ctgtctctac	taaaaaatac	48180
aaaaaagaaa	gaaaaaaaaa	ttagccaggc	atggtggcgg	gctctttag	tcccagttac	48240
ttgggaggct	gaggcaagag	aatggtgtga	acccgggagg	cagagcttgc	agtgaagcca	48300
gatcgacca	ttgactcca	gcctggcgga	cagagtaaga	ctctgtctca	aaaaaaaaaa	48360
ggaaaaagaa	aaaatatata	tacattgtgt	actttttggc	atctggttta	ttttgctcaa	48420
tatcacatct	gcgaaattaa	tctacactgt	gtgtatgaaa	ggttggttct	ttttgttgtg	48480
atgcagtatt	ccgtcgtgtg	actacgggac	aatttgctta	tccgtattcc	tatcggtggg	48540
catttggtgt	gttaccaggt	tctggctgtt	atgaataaag	ttgctatgga	tattcttgta	48600
cactacttct	ggtgagcgta	tgactcatt	tcgcttatgt	aaatatcttg	ggtggaatta	48660
cctgatcata	aggtaggtgt	gttggttttg	taatgtgtgt	acttggttat	gctgaattcc	48720
cttttttgtg	tatttctggt	tagagcgga	catgagggtg	tctcttcagg	gaatctggag	48780
ggtggaagg	aagcaggagt	cggtttctgg	ctcacacatg	ttgtgactga	actgctggta	48840
cacctgggtg	gcatggagct	ggcttctcct	ttggcggtgc	ctactgttgg	ggcagggtgtg	48900
tatgtggtta	gctccatgca	atgaacccgg	gcttctgcaa	aatacattaa	caacgacaga	48960
gacaacaaaa	gctgatgtgg	atttaaaggc	ttcagttcan	nnnnnnnnnn	nnnnnnnnnn	49020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49080
nnnnnnnnnn	nnnnnnnnnn	cagcagtgtg	tctcaactga	gcatagtttt	gcctcagagg	49140
ggacatttgg	taatgtctgc	agacattttt	tgattgtcac	agcccagccg	agaaggtagt	49200
actagtatct	ttttggtaga	ggctagagag	gctgctaaac	atctaacaat	gcacaggaca	49260
ggcctctgta	acaaaaaagt	atccagtcac	aatgtccac	agtgttgaga	ggtttaggta	49320
agtaggcgct	aaaacataag	gagactgtgc	ctgagagcaa	gaaggagtaa	ttggaaagt	49380
ctggtgtgat	tagctctggg	ttttagaaag	ctcatttttg	ctgctttag	acagtgcac	49440
agaggtggag	gaggtggga	agactggag	caggaaagt	aatttgggag	ccactgaaat	49500
gatccagggtg	aaaaacggtc	agcaggtgac	taggaaagt	gcagaggcaa	tggggatggg	49560
tggctggatg	agatggtgaa	gaaagcacta	taactaacta	atgtgtggat	gatgggcagg	49620
aggggtgaag	gatgaccaga	gtcctgcctt	gcaggtctag	ttggaagggt	atggtttctc	49680
ctgagaaagt	gaccacaaaa	agtgaagcag	gtttgtgcgt	gtgtgtgtgt	gtgtgtgtgt	49740
gtgtgtgttg	agttcagctc	gagatgtgtt	ggactcacaa	tgtccatggg	acatccaagt	49800
ggagaagcat	cttgggtgac	catatgtgtg	agtctgcagc	tcagaaacag	gcctggggct	49860
ggagatgaag	acttgggaat	gatctgcgta	tatatgttgt	agcttgagcc	acaagagtag	49920
atgacataac	ccgtggtggg	tgtgcagaat	taggagagac	gtgcaccaag	aagccagggt	49980
atccccaata	tttaaccatc	tggaagaata	agaggagcct	gccaacagaa	attgggaggg	50040
aatggccaca	aaggctactg	agaagggaag	cagttcttaa	gaagggggaa	gtgaagaggt	50100
atcactactg	cagaggctca	gtaggataag	aactgaagaa	tgtctgttgg	gtttggcaat	50160
ggggtagtca	gtgggcacct	gggcaaaagc	agttttgggt	gagcaatagg	gataacagaa	50220
acaagactgc	tatggtaaga	ggaggaagag	ggtgttgagg	aagtggccag	cgagtctaca	50280
ccacttgctg	gaggagcttg	gctttggtgc	agccagctca	ctcattgact	ctcattgact	50340
taacctccaa	gaaacacaaa	atcatccata	tcctggctca	aattccagca	ctaccaggag	50400
atgggtggcc	cctagaaatg	ccatcccact	tctcctctgc	ttatcctatc	ctatctgtca	50460
gtctgttgag	cccaggctaa	gcgctacctc	ctcaagcaag	ccttctctgc	ctgcggtcac	50520
actttaagt	atcctgacaa	cactgaaaat	gtgtgtctct	tccattcatg	ttagttctac	50580
acttctgagt	atctcctcaa	tatattgcct	tgttttacta	atatgctcgt	tctgtttgcc	50640
ttatttatca	gctaccttaa	acctcctgc	aactagagat	tctctttaag	tatttgttga	50700
ataaatgaat	gaatcaatcg	atgatccaga	gcctgtgtgc	ggctgtgtgc	catggtggat	50760
gaggctcaga	aaatacctgt	agaatcgaaa	ttaatgcattg	tgtgctctga	tctaaactca	50820

gctaaacttt	ctccaggggg	ttaaagttaa	gttgattagt	caattgatta	attaattcat	50880
tatgtaattg	aaaaactcct	tctatgacct	gggcagagtt	ataggcagtg	aacaagacag	50940
acaaggtcct	tgttgtcatg	aagtttgctt	tctgaaggag	agagataata	aacaagaaac	51000
cagtaagaaa	gcaagattat	atcatttttg	taaatgttct	tgtggaaata	aatgtgatga	51060
tgtgtaacaa	aagtacaaaa	taggagagtg	gggtgggttg	gcttctttta	gaaagagttc	51120
tcggagaagg	cttatctgag	gaggtggcct	tttaaccagt	acaaatgctt	tagcttggcc	51180
agtggagctg	ggaccaggat	gacaagggct	acttgtcatg	ccagtgagtt	tgagcttgta	51240
gacaagagcc	tgatcatgaa	agactttgca	gatgggtgta	atgggttttg	gttaattgct	51300
actatgtggg	aagactttga	atgggaagca	tggggacaat	ggcctgtgat	acatgttatc	51360
aaatatgggt	gcaggggcta	gtgagtgggc	agcagagata	gggagaagta	gacggactgg	51420
ggaaggtaga	agatggggca	ggggaggcaa	ttactgcaaa	gacataattc	ttctaagctc	51480
actgagtggt	catggtctct	gggagcagag	gttctctggg	gggaaagagg	ataatgtcac	51540
ttcctgagga	agcgggaaga	acccatctga	gacgtgggga	ctgtgctggt	tcgtttctaa	51600
ggggccttcc	agatctcaca	tgccaatcgt	cttgggtctat	gtcaattggt	ggggcatcca	51660
aatggggaac	tggtgtccag	gccgatttca	cagaacaacc	gcccgagtcca	tatctcccga	51720
gccattcacc	cttgacagtg	cgtagctctc	ttcaccagct	tttatctgct	ccgtggggat	51780
gttgggcctt	cccaggttaac	aagcagttga	tcagcccgag	agatcaggtc	cctggagctc	51840
gtcacttttc	tgaggggtgg	gagagaatcc	tggagcagaa	catgtaacta	gaaggggcac	51900
ctggcttcc	atgggtctgag	ggagagaatg	gtgggatctc	tgccctgaat	caaacctccc	51960
tttctcagtg	tccatcttac	ctctctgctg	taccttcgtt	attttccagc	agctcctcag	52020
cccgttccct	tgggaccctt	ctctgccaat	ccctacaccc	actgtaaatt	tcaccgtggg	52080
aggagagatg	gccttgaggg	ctgtattagt	cttctattct	gcataacaaa	ttgcctcaaa	52140
tttagcagct	tcaaaacaact	catgtttatt	agctcatcgt	gagttcatca	gcagtggtgg	52200
cccagcatgg	catggttttc	tgctcagggt	ctcacaggc	taaaatcaag	atgtgtctct	52260
ggctgtgtgt	tcatctggag	tttagggttc	tcttccaggc	tcacgtgggt	gtggcagaat	52320
tctgttccct	ggagttgcag	ggctgaggct	ctgttttctt	gctgactgtc	agatgagggc	52380
tgctctcagg	tcctcgaggc	tgcccacatt	gcttgccacg	tgctgtgtct	tttccatcct	52440
tgaagccagt	gatggagaat	ttcccttgga	ttgaatcacc	cacatgggtg	gactctctga	52500
cttcaggaag	agagccctgt	ctcttttatg	ggatcacctg	attagatcat	acccatagag	52560
ggcagttcct	tttccctaaa	gtcaactgtg	gcattgtaaca	tcacacaacc	acaggagtaa	52620
aatccatcat	atttacagtc	ccagggatta	tgcacagtgc	accaggggac	aactgaattc	52680
tgctgtcaaa	aagggccaaag	caggacttta	ttggtgaaga	acagtgggaat	gtcattcttg	52740
gttcttccag	aaaaaaatca	ctcagtaaag	ttagagggtt	tcttgccctt	tggggaagtca	52800
tcaaagaatc	tcatggaggg	tttgacattt	caccctagaa	acatcacacc	atgttttcta	52860
taattgcagg	gttcattggtc	ccttgaagcc	tattcatagt	ttccaggttg	aaaagctctg	52920
ctgcaggggt	tggggaggga	tgacaggtga	ggtagagggt	gaatagtgtg	agctgcata	52980
ctggagctgt	ggtggttttt	ttagtcttta	agctgtcatg	tgttgggggt	tgggcatggg	53040
aggggcatcc	caagagctcc	ttggtattga	caccatctcc	aaggtgatct	ctgctctgcc	53100
tggtgcacac	atgtttttct	cctgttgcaa	cagcccactc	ttgtagaaga	gcagacccct	53160
cagtaccagg	tctgaccctg	gacagcttgt	accaggagct	acagcacact	ccccacaaag	53220
cctaaagttg	ggatgagccc	cccagaaatt	agatcagaaa	agattaaatg	cagagggtgat	53280
ctgtcagggt	ccctttggaa	gtgctggtat	ggagaggatt	gactgagctc	gttttaggaac	53340
ctccaagctc	tgtagtaact	ttagggctag	aaaggaggat	gcctaagatt	caggatcctg	53400
cagtgatgag	tcaacatttc	ttggggaagg	agccagggct	gaggattaaa	cggagatgat	53460
gggtatcggt	ctcttgctca	aaggcactgg	accccaaggc	ctccagctct	tcgtctccat	53520
ttgaaattca	agtccctgagc	acaccacagt	tgtgatgcag	ggaaagaatg	tgcttatcag	53580
agagcctggg	caagtggggc	ccttgtgagt	accgttcaac	ctcatttatg	tcattggcac	53640
caaaagtaga	catcagctct	ttgaaagttt	gattaatgct	ggtcacactc	aaagaccctg	53700
ggtagcatte	atttactaag	caattactaa	ataccagttt	ctgtgctaaa	tgctgcatca	53760
gtcagggtcc	ttaatggcag	gcagcagaaa	ctctccttgg	ctgatctaag	tagaaaaatc	53820
caggactgaa	aggaacggga	gtagctcatg	aaattgcagg	aaggggccgga	aaaccagaca	53880
tggagccaaa	gtcaggctgc	agaacaggct	tagggaggat	cccactgctg	ctgagacctc	53940
gaccttgtgt	ctggcaccca	ggatgttgta	gggctcagac	cctggatcaa	tgtatcctgc	54000
agtgcctctg	tgggtactgc	aactccagga	actcaatctt	gtcaacgcca	ccgccagaga	54060
gaggccttct	tggcctccat	cttttttggt	actagctcca	gattcaaaat	cttgaataga	54120
tgcttcttct	ctttgataga	gcccagtcct	atgctgttag	tgcaaaggaa	gctgaaaatc	54180
tattaggaac	ttttgtcttc	aaaaatgaga	ggcctgtcct	ccaccaagat	ccataggaaa	54240
tggaaatccaa	gaaccacag	gaaggggtga	ggtagctggg	cagctcacag	catgcatgct	54300
acatgtgaat	tatctcattc	atctctcaca	ctaccagtg	aggtaggtat	tgtcatccct	54360
acttcataaa	tgatgatatg	aggtacagaa	agtttaagga	acttgcccag	gacacgacac	54420
gcagctatta	agtgctagac	ccagtcgaat	tgagtctgac	ttggactgtc	tgactccaga	54480
agccaccctc	tcagacactg	ctgtatactt	ccagtgaatg	ttgatgaaat	tttcagggtt	54540
gctaagctgt	ggatttcaga	tcctggattg	tatgacctaa	aagagagact	tccttaggag	54600
tgagggtccc	tgaacagtca	actgggttcc	aagaatgggc	tcctctcat	caccttatga	54660

cagtaatcct	ctgtccaaca	gccaaagagg	tcctgtgggg	agggcttgca	gatgggagtg	54720
cgcagagccc	agctcaaagc	tcctgactag	gctcttggtg	agtattcctt	tgattcctgc	54780
ttctgtcttt	ttaaatacaat	ggagacaggg	gagggttatc	tccatcctcg	gctcaagatg	54840
aaatgcatcg	ttcctcgttt	ttctcattcc	ttcccaatgt	gtgtactgtt	aacttttagtt	54900
atgaaggaaa	ttacagtgtc	ctgtgcatat	accaaggctg	tccaacctcc	acacctttgc	54960
tcaagctgtt	ccttctactt	gaaatgcctg	tttccctccc	ttctaattgc	atctttccat	55020
ccaggtagga	atcagctcct	tggttcatgg	agccttttct	gctctgtttt	actatgcatg	55080
gacttccttc	tgaattagca	gaggatgttt	cctagcttgg	tcttaaccct	tctccttttg	55140
tttgacctca	atcttactcat	cttacaaatt	agggtgtaag	ctaattgaat	acaggatcta	55200
tgcttcactc	tgatttttatc	tccacctgga	tagcatcatt	tttgacacac	aagcaggcat	55260
atgggagggg	agagaagttt	ggtgccagaa	agaactggat	ttgaattcta	accctgttgt	55320
ttacgtgagt	acgttactta	accattaatt	acttcaatgt	atattttatta	agtacctact	55380
atgtgccggg	cactgtacta	agcaccaagg	atacaatggg	gagtaaagag	atgcagcctt	55440
caccatcacg	aaggaagaca	gatgttaatc	cattaacca	gtaatctcac	aagaaaagta	55500
aaatgactaa	ctgataagga	caagcccctg	gagctacaag	agggtgtata	cagggcatcg	55560
atccaataag	ggcagtgttg	cggggagatc	aggagccaca	cagagcctgg	gttgtctcac	55620
ttgaaaaatg	gggtatcaac	cacctacctc	actaggtttt	taaaatcagg	ttaaatgagg	55680
taataacttg	catgaacagt	atcttctgtg	ttgatgattg	attgaaacgg	agtctcactc	55740
tctcgcccaa	gctggagtgc	agtgggtgca	tctcagctca	ctgcaacctc	tacttctctg	55800
gttcaagtga	ttctcctgcc	tcagactccc	aagtagctgg	gattacaggc	agccaccctt	55860
atgcctgact	aatttttgtg	tttttagtag	agacaaggct	ttgcaatgtt	gaccaggctg	55920
gtctcaacct	cctgacctca	aaagatccac	ccacctcagc	ctcccaaagt	gctgggatca	55980
caggcatcag	ccactgcatc	cagccacttg	ccatgcattg	catttaaaaa	tgttcagtaa	56040
atgttaccat	aatgaaggct	ggtaggttgg	ccaactgagt	ggtctgattc	agaaggaaag	56100
aagttagaca	tacgtgaaca	tttctgttac	ttgaagatcc	tcaggacagt	gactcctaga	56160
cccattcttc	atcacagtca	gctgggaagc	ttttaaaaaa	atgcagacat	ctgaccttca	56220
cgctagacct	attagccaag	cagaagtttc	tgggcagggc	atctgcatat	ttttaaaaat	56280
ctttaataag	gcagcctcaa	aattacagat	tcagcacgca	tttaccataa	ccactgaaga	56340
aatgcaaagt	tataaaaaaga	agataaaaca	caatctgtct	cctgctttct	tccctctcct	56400
cccctgcttc	tggaggcaac	aagggtcaat	atttgggtgtg	attcctttta	gcattccctc	56460
catcaatggt	cacataagga	tgctcacaga	taagcaccta	tgccgggggtt	ttttttttcc	56520
ttgtaaaact	attcacatac	ttaaatacttt	cctcagtatc	ttgccttttt	tcacttcatg	56580
tcacagaaac	atctcttcag	gtttatagat	acagggtccag	ctcttctttt	catagccata	56640
taacattctg	tagaatagag	aggacacatt	ttactcagtg	tccgattgat	ggatatcaat	56700
attgttttca	tttctacaaa	tagtcaagga	ataacataac	tctgtaaaag	ttttattact	56760
tataggcgca	tttatgccta	aaggatagtc	tcaaaagagt	gaaactgatc	aaatgtgcat	56820
ttttttattt	taataagtat	ggacagattt	gtttctaaaa	tgtttgtggc	agttcaaaac	56880
accagtaaaa	cagggggagat	atgtattttg	gaaaagcacc	caaggcgatt	ctgaagtgtg	56940
gcccaggata	agaaccattg	cccagagctg	ttccagatgg	cccctgggtt	cctgaagtgg	57000
gtatcgggag	agaaatcttc	actgaatgaa	tgagtgggct	ccccagggaa	gtgatgaaat	57060
ggtcctttatc	agccttgcta	tctccctctg	acagaggcaa	actctctctc	cctgggggaa	57120
gttccctcaa	ggcctctata	taagaagtct	ttgtgagagg	aagcaaagaa	ggacctgggc	57180
tttgggaagt	tctaaagacc	caggaaggte	tctgggtggg	tgagtgcctt	ctctgtctgtg	57240
gtggagctgg	tgacagttta	ttctcccagg	aggtccctgg	ctgtggctga	cagtttcttg	57300
agggctggca	ggcgtctacc	tgtggctttc	aggttatgag	gatgtcagca	ggggcagcct	57360
tcatectctg	ccttgacat	tccttctgctg	ggatgtgaaa	gtgctccttg	gctggggaaa	57420
ggagatggtg	gagacatgga	ggagggtgtg	ggtggcttct	tgaactctga	ggaggggaca	57480
taccttctaa	gtcctatgtg	ttcctaggaa	agccaataat	cattgcttct	cccgcctttt	57540
ttatgtcata	gactctgagg	gaccatttaa	gtacaaaca	ataagcgtaa	tagtcccttc	57600
tttacttccg	ggcctgaagg	aaagccagcc	tcagccaccc	ctcagggttt	gctgcgttct	57660
gtttagaaag	aggctccttg	gtcctggatc	ctggagcatc	aggagctggg	cctggcatga	57720
gcttttctgg	cccattcctga	tttctattca	ggccttcttt	ttctccacct	cactcccacg	57780
gtcccttaat	ggtgtgattg	tgatgtgtgt	gcatgtgtgt	ctgtgtgtgt	caatgacaaa	57840
ctgtgttctc	cgttgcagga	taaagccaag	atgaaactcc	ccttacttct	ggctcttcta	57900
tttggggcag	tttctgctct	tcattctaagt	aagtgttttt	tgcccttcagt	ctttctttct	57960
ctgttttttc	cctttctatg	gtagatgggg	tcagagttac	acacccaccc	ccttctttga	58020
tcgtcttcta	tttctgaatt	tctgtgtgct	tcaagggtatg	gggactctat	ggccaggagt	58080
tgaaaggatt	tctcaaggcg	tctgttatgt	ctgtggtctt	ggttctactg	tgacattccc	58140
aattttgtcc	tttctccatt	atgcttactt	tgagcttact	gagtgccttc	tctcctttta	58200
ctctcttagc	atcgccatga	agtaggtggt	attgtatacc	catttcacag	aaatacagct	58260
ggtggatgat	ggaaccagta	cccaagccca	tgactgcccg	actctaagtc	catgctctta	58320
accaccttga	ccttgtcagg	cagcttgggt	tccctccta	gagactgggt	tccaggttcc	58380
ccttcccagg	cagagttgag	cactctgatg	cccagggcaa	ggtgtgagct	gtctgtgggt	58440
ctggggagga	acaaggggag	atgtgaagga	aggacactta	gctatcctcc	ctgccagggt	58500

ctgagacttc	caccttttgg	acccttttgg	gtgctaagac	gctgcctgag	gatgaggaga	58560
caccagagca	ggagatggag	gagaccctt	gcagggagct	ggaggaagag	gaggagtggg	58620
gctctggaag	tgaagatgcc	tccaagaaag	atggggctgt	tgagtctatc	tcagtgccag	58680
atatggtgga	caaaaacctt	acgtgtcctg	aggaagagga	cacagtaaaa	gtgggtgggca	58740
tccttggttg	ccagacctgc	cgctacctcc	tggtgagaag	tcttcagacg	tttagtcaag	58800
cttggttgag	tgccctatgg	ctgaggctga	ggtgggagca	tggaacgggt	gtgggatatg	58860
ccccagcat	tgctatcact	ggctcttttt	cccattgagg	gccctggggg	tgctcagtaga	58920
acctgagcct	cagagaggtg	ttggggtaag	aggggagggc	cacctacaaa	cagaagttgc	58980
attttgggtc	ccaaccttca	aatgggttgg	gcagggaggg	gagggaaatga	attgtgggga	59040
ctcaagaccc	atgtgaattc	atgtaggaag	gatgctccat	tctttgtctt	ttatcctgcc	59100
ctgtagttta	cttgccggag	gtgctacagg	ggcaacctgg	tttccatcca	caacttcaat	59160
attaattatc	gaatccagtg	ttctgtcagc	gcgctcaacc	agggctcaagt	ctggattgga	59220
ggcaggatca	caggctcggg	aagagaagtg	tgaacactaa	atggggtgca	cctgctgatc	59280
tcagccagca	ctcagcttgc	atcagatttg	tctgtttttc	tcctgtataa	tctccagaag	59340
aaccagggat	agatggacac	ccacagacaa	cactgagggg	gctgcctggg	cattcaggga	59400
agagctaagg	atttagaatc	aggaggtttg	gggtccaagt	cctttccatc	tctcactatc	59460
tatgtaactt	aagtttagctg	ggcatgggtg	tgcatgtctg	taatcctagc	tacttgggag	59520
gctgaggcag	gagagtcact	ggaacctggg	agacagaggt	tgccgtgagc	cgagatggag	59580
ccattgcact	ccagcctggg	caacaagagc	gaaactccgc	ctcaaaaata	aatataataa	59640
taaataaaat	aaaaaaaaaa	ttaaaacaag	accatgagtt	tgtttccctc	tctctaggat	59700
gagtttgcaa	cccttgttct	acctttttgt	agggctggaa	ggacaagcct	gtcactggga	59760
tgcatagaat	ctgatgggtg	taattggcgt	ggatcagcat	ttcagatgac	taggacagtt	59820
cccatcatgt	tccagcaggg	aagggcccat	tgcccgttgg	gcagcagaaa	gagctggcag	59880
atacggggcc	aggtctgctt	ctctgccttc	cctctgcccc	atcccttctt	ccccctctgc	59940
tttctccagg	gtcgtctgag	acgttttcag	tggttgacg	gcagccgctg	gaactttgca	60000
tactgggctg	ctcaccagcc	ctggtcccgc	ggtggtcact	gcgtggccct	gtgtacccga	60060
ggtgaggtgg	ggctggggat	gaacgatgga	aaggtctggg	agatgggaag	tgcccccaagg	60120
aggagatgct	acaaagagcc	tgaccctttg	tgggagaggc	ttcctgggtc	ttttatatac	60180
tctgactcca	cagcagtgtg	tgggtgggaa	aagaggccct	cctgtgggtt	gagttgggat	60240
ggacaagagg	ctgaaaagtc	ctttctgttc	tgcttccaca	ggaggccact	ggcgctgagc	60300
ccactgcctc	atgaagacttc	ctttcatctg	ttcctactga	gctgggtcca	gccagcagtt	60360
cagagctgcc	ctctcctggg	cagctgcctc	ccctcctctg	cttgccatcc	ctccctccac	60420
ctccctgcaa	taaaatgggt	tttactgaaa	tggatttatt	ttctcctctg	atcgcggtac	60480
cactctgctt	agccctcatt	gaaacttctt	ccttatcatc	tctccccaca	ccacaacttt	60540
catagaagtg	tcagaagcta	ctactccttg	aggaggagga	tggaggggtg	agttgggtct	60600
atggagcctt	ttggagatgg	aggaatgggc	tcagctagtt	ctcttcatag	aacacttgat	60660
tactgggcac	ctgcatagtg	ctgccaggac	ctttcaaggt	tgtaggtaga	ctcccaatgg	60720
cccagtttgc	atctctgtaa	ccaaaggcct	tttctctctc	tctctccaac	cccagaactg	60780
tggttggttt	tatatgtaag	gaagttaaca	tgctccctgg	aacagtccac	aacattcagg	60840
aatgaatgta	taagtaccgc	aatccccggc	ccctcaagtg	gaataaatct	aacatgtatt	60900
gggcaccatt	tcccagtggc	ctgctgtggt	agttggcctt	attccatgca	tttttatggg	60960
ctgccttccc	ttcctcaact	gcattctctg	ctccttcta	ctctctgcaa	ctcccaata	61020
aacacttgta	cgcaactccc	tctctcagga	tctccttctg	gggaaacctg	atataagaca	61080
gcttgccatg	cgtcagactc	tgaatgaggc	ctgggaatac	aagacatagt	cctctggcac	61140
ttgggatata	tggttatttg	taacataggc	acaaaaacat	ctactagttg	ttatcgctta	61200
ttgagcacc	acaacatacc	ccctgctgtg	gcaggcacct	tgccatagatg	acctcatgtg	61260
atcaataatt	atgagcccta	ttttacagaa	ccaggctcag	agaagttagg	atctgtcaaa	61320
agacttgccc	aagactgaac	ctctaaatgc	aactcatatt	gaaattcaac	tctgctccaa	61380
agcatgttac	tttaaccctt	gtgcttttac	agctggctac	tctcccttta	tggtcacacg	61440
gggatgaagc	acggggggag	gaaagccaga	ctgtctcact	cctgggttca	tcttgggaca	61500
caggacacca	gcccagctgg	aggtgagggg	gctttaatca	gaggggaggg	aggaaggcat	61560
tctcaacccc	ttctgtacta	gggaggtcag	cagaagaaaa	taattcaatg	ttctaaagcc	61620
atTTTTTTTct	ccagcattcc	tccaattcat	agatcttcat	atgggattag	gggtcagag	61680
aggggtgaaa	caagaactct	atTTTTTTTgg	agtgtggtat	agagaaggga	tgctacttct	61740
ctaaggtcac	atagtaagtt	gagaaagaga	gagaaatcaa	actcaggttc	atttcaacta	61800
ttgttccaca	agaatctgtt	gatttcaaaa	atggtggact	atgggttcat	ccctgtgggtg	61860
agtgtgtgta	ggaatgcagct	gaggtggaac	tttcaactct	tgccctcttg	gactttatat	61920
tctggtgtgg	aaaggcattg	cttcccttat	ttcaatatta	acaacaaagg	gtaataatat	61980
ttcccattta	tttaagcattt	actaggtgtc	aggtactgtg	ctaaatgtta	ggtgaacttt	62040
gtcttgttcc	tcataaatct	ctgccgctgt	gggtgtgtac	tttgacagaa	gtttgacttc	62100
cagtccacag	agatcttctt	tgggggagta	atatcaagaa	ggggcacgaa	ggaagctgca	62160
gggctcctag	tcccatcctg	tatctcgacc	taggcattgt	tacattgggtg	cattcactgt	62220
gaagtttccc	tgagcagtc	actctatagt	gtgctttata	ggagcacatt	gtacatccat	62280
tgaaaaattt	ttcttggccg	ggcacggtgg	ctcatgtctg	taatccagc	actttgggag	62340

gccgagacag	gcggtatcacc	tgaggtcggg	agtttgagac	ctgcctgacc	aacatggaga	62400
aaccccgctct	ctactaaaaa	tacaaaaaaa	ttagccgggt	gtggtggcac	atgcctgtaa	62460
tcccagctac	taggaggtt	gaggctggag	aatcgcttga	acctgggagg	cgaagggtgc	62520
agtgagccga	gacgtgccca	ttgcactcca	gcctgggcaa	caagagcgaa	actccgtctc	62580
aaaagaaaga	aagagatttt	ttctttttct	taaaaagtaa	aaatcatgaa	ataaggggac	62640
tgggctaata	ttccaaaata	tgggtttgtg	tgtgaatttt	cctctccagt	aagatactaa	62700
ctaagctctg	tgaactgtt	tatctatggt	tctttatcat	tgaatccttg	gagttccctta	62760
cactgtgcag	agcacagagt	aggggctcaa	tcaacagtgc	actcattgct	ttttcataga	62820
caagggccac	cctcactcaa	ctcatgtgcc	aggcatagtt	ctgagagctt	tgcttaagct	62880
gatnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnggtcact	aatttgggtat	63000
agagtattgt	tcattgattt	cattttattt	tgtttctgat	ttttaagat	tgttttactt	63060
gttttcttcc	tattattatt	ttattttatt	tgtaaaacat	ttacatatca	gacatttaca	63120
ttttcccaaa	ggtaaaactg	tgaacaaga	tatattcaaa	gaagtttact	ttccctctct	63180
gtttcttgta	ccccctttcc	tcttctttag	gtaaccattt	ttattttttt	aaatataaac	63240
attgtgtagg	tgtatatata	tgtattagtc	tgttttcatg	ctgctgataa	agacctatct	63300
gagactggga	agaaaaagag	gtttaattgg	acttacagtt	ccacatggct	ggcaaggcct	63360
cagaatcatg	gcaggaggtg	aaaggcactt	cttacacggt	ggtggcaaga	gaaaaatgag	63420
gaagaatcaa	aagtggaaac	ccctgataaa	cccatcagat	ctcgtgagat	ttattcacta	63480
tcacaagaat	agcgtgggaa	agactggccc	ccatgattca	gttaccctcc	cccactgggt	63540
cccacccaca	atacgtggga	attctgggag	atataattca	acgtgagatt	tgggtgggga	63600
cacagccaaa	ccatatcaat	acatttccct	ctcttttaga	taaaaggtag	tatactgtat	63660
acataattct	gcagagtttt	tttttttttt	gatgtaactc	tatcctgagg	gtgctctgta	63720
gcaggacct	ctatgcctt	ttaaccactg	cctgggtctc	cattacatgg	ctgcagcata	63780
gttgccacag	cattcctgta	ctgatgacta	tttgattgt	ttccagtctt	ttgctattac	63840
cagtagtgtt	acaaagagga	tctggctaca	tggtcagggt	ggggaggggc	agatgtgtag	63900
cctgtcagga	gggtattgca	gtaatccatg	actgagttaa	tggtagttaa	aagctaggat	63960
gagtcagtgg	ggttggagag	aagtgggcac	atttgaatga	tatgtaggag	gtgaatgatc	64020
agcattattg	atgagtttga	ggtggggcat	gtggggaaag	gattcgagga	tgactcccag	64080
gtttctgttg	ggacagtggg	tggatagtgg	ctcctcccc	ttttccaatc	ttccttgccc	64140
cttcgctgac	tctctgttgg	tggcctaca	gagagcttct	ttttcctctc	tgttcgccc	64200
ggttccctcca	ctttggcggt	ggccctctgc	tcgacgggtg	cttcgctggc	cctgacatcc	64260
ctgctgtgcc	tgggcttcgc	cctctgtgcc	tcagtcccca	tcctccctct	ccagtacctc	64320
accttcatcc	tgcaagtgat	cagccgctcc	ttcctctatg	ggagcaacgc	ggccttcctc	64380
acccttgctg	aagtggcctt	ggggcgggct	ctgtggagac	ggacacactg	gggcaaagag	64440
aagctggagg	taaaagaaat	gggaggcaag	gcggggcctg	gaggcagtca	ggtgcgggag	64500
actgggtttg	ggggcagggt	tggagggggt	gagaccagag	gtggtgggaa	ggatagaaca	64560
ttcatgcact	tgagccttta	catctgcggt	gccctctccc	tctgttttct	acctggtgaa	64620
ctcgtattca	tcctctgagg	cccacttctg	tttcagttct	ccagggaaga	aatggaaaag	64680
tgtcttccct	tctttgtgcc	cttagtactc	tagtcttact	tcctttgcta	gtgcgtgcat	64740
tgtctggcat	gccatccatt	tacatgcctg	tcttttcttt	cctgggtgcag	cctgcatgag	64800
ggtcctgtct	gtttttccag	ggccccgcct	gtgccttctt	ctgggttctg	tgggtcaa	64860
gtctgagcag	agctgaagag	ggaaaggcca	gacaggtgtg	gttggagggc	aggcctagga	64920
caggggagct	ggggacaagc	ggccgacagc	ccccagggc	caggcttctg	cttggaggga	64980
gggtccctga	agctcactgg	aacccctctg	gtttctctcc	ccagtttccc	ttcagagcac	65040
tttggaagc	tctttgggct	ggtgatggcc	ttgtcggctg	tgggtgtctct	gctccagttc	65100
cccatcttca	ccctcatcaa	aggctccctt	cagaatgacc	cattttacgt	gagtactggg	65160
aggatgggga	tccttgccag	gaggcctggg	ccttaggcct	tggtgcctcc	aaatctggct	65220
gtgatggcct	gggtatgtag	catggtgcag	cttcccaaag	ggtctgtgtt	attcaagtat	65280
ttggggcaaa	agtattttgt	tgtgtgggga	aacagacatt	ctggagtagg	gtgggggaatt	65340
ctcacgaaac	ttcaagcaaa	atcctgagac	ctcaaagggt	tttccctgctt	gtggtgagtg	65400
caggccacc	ctggcctctc	ccctaggccc	acacagggtt	tcacagttg	gccccaggga	65460
caggacctct	gtgctttcac	ctctgtgtcc	ttacacctgg	agggatgctc	tgaggtcctg	65520
ctctaggagg	tggtcgtgag	tctcctgctc	tttgagaaa	ctgaggctca	aagaggttac	65580
ttacgtgttc	agaggcacca	gctaaggagc	aaaagtcaac	tttgaattct	gtgttttgac	65640
tactgacag	ctctattttg	ctcatttttt	atttttaag	cagcaaatct	tagaatagga	65700
gttttaaatc	atcacttggg	gaaaagaaag	actaaatgtt	ttttgttttt	gttttgaga	65760
cacgatcttg	ctttgtcacc	caggctggag	tgcagtggca	caatctcggc	tcactgcagc	65820
ctcgatctcc	tggactcaag	cgatcctctc	atctcagcct	cctgagtagc	tgacactaca	65880
ggcatgtgcc	accatgccaa	gcttattttt	ttttattttt	ttgatagaca	ctgggggttc	65940
gctatgttgc	ctgggctggt	tttgaattcc	tggcctcaag	cgatccaccc	gtctctgcct	66000
tccaaaatgc	tgtgattaca	ggcgtgaacc	actgtgcata	gccaaaagag	taaaattgaa	66060
atctgaggcg	aatgacttga	tgtgacatc	aggtagccta	gtaatcagct	gtgtattcta	66120
gctggtgcct	ctaccagctt	cccatgtgac	cttgaacatg	tcattgaatg	ctcgctaggc	66180

ctctgtttct	ttatctgtga	aatgggcttg	atattcctcc	tctaccccaa	ccgatagtgc	66240
agaatgaaaa	gtaactgaaa	gtccttcctc	cagggcacca	tagtgtctgg	gtgaaaagta	66300
gaatataaac	tcggtagact	tctggccctc	tcattgggtca	tggaaatggac	cagtgtctgc	66360
ttcattgagc	aacagttctg	ttgttcagaa	ttcctggatt	tcacctcact	tctgtctctcc	66420
ctgcaggtga	atgtgatgtt	catgcttgcc	attctttctga	cattcttcca	cccctttctg	66480
gtatatcggg	aatgccgtac	ttggaaagaa	agtccctctg	caattgcata	gttcagaagc	66540
cctcactttt	cagccccgag	gatgggtttg	ttcatcttcc	accacctttg	aggacctcgt	66600
gtcccaaaaag	actttgccta	tcccagcaaa	acacacacac	acacacacac	acacacaaaa	66660
taaagacaca	caaggacgtc	tgcgagcaaa	gaaaagaatc	tcagttgccca	agcagattga	66720
tatcacacag	actcaaagca	aaggcatgtg	gaacttcttt	atttcaaaaac	agaagtgtct	66780
ccttgacatt	agccttggca	gacccttgac	tccaggggag	atgacctggg	ggagggaagt	66840
tgtcaactat	ttcttttagc	ctgtttggct	ccgaagccta	tatgtgcctg	gatcctctgc	66900
cacgggttaa	attttcaggt	gaagagttag	gttgtcatgg	cctcagctat	gcttctctggc	66960
tctccctcaa	gagtgcagcc	ttggctagag	aactcacagc	tctgggaaaa	agaggagcag	67020
acagggttcc	ctgggccag	tctcagccca	gccactgatg	ctggatgacc	ttggcctgac	67080
cctgggtctg	tctcagaatc	acttttccca	tctgtaaaat	tgagatgaat	tttgggtgtt	67140
aaagtctctc	ctggagcaga	tgtcctagaa	ggttttagga	atagtgcag	agtcaggcca	67200
ccccaaaggg	catgggagcc	agctgacctg	cttgaccgaa	ggatttctga	cagactatct	67260
ttggggatgt	tttcaagaag	ggatataagt	tatttacttt	gggcatttaa	aagaaaattt	67320
ctctcgga	taattttata	gaaaaataaa	gcttctgtgt	ctaaggcaac	tactgtttcc	67380
atctctctag	gctttgggcc	ggggctgtgt	gtgtgtgtgt	gtgtgtgttt	gtgtgtatgt	67440
gtatgtttct	gaggaggccc	taccctggca	tgagagggta	gggaatctgg	ctacacatct	67500
agtgtggcag	ctggaccag	aggtggggca	ggaacctga	ctatgattca	ccccgctggt	67560
cctgggatgt	gggccagag	acttctctcc	ccaggaaacc	ctctgcttcc	tcttctctcc	67620
cacatcctta	actaacttta	gcagaacctt	actctcact	acacaccccc	agctagaagc	67680
gctggatgga	atcagaaatt	cctagtttga	gtttcaattc	tgccccctcag	cagctgggca	67740
agccccctaa	ccactctgag	tactagtctc	cccacctgca	aagtgcagtt	aatcatttct	67800
atctctgatg	gcgattgtga	gaatgtaaag	tcattgcaac	tgccctagcac	atggtaggag	67860
cacatgaggg	tttgctcctg	tgtttactca	tgaccttgg	ggaggacggg	ggcaaagagg	67920
gagaagttga	gggtgcagga	ggagagatgg	caggtgggtg	ggatgggaga	atctggggca	67980
cacctgctgt	ctcattccca	ccttgctagg	agagggaata	ggaaagaaca	gtgggagcca	68040
gggggatggg	ggtggaaggc	agggggtggc	agggcaggtc	atccatccat	tcattcaaca	68100
aatgtttatt	gagcacctgc	cacgtgtcag	gccctgtcct	gggtgctggg	gctataaaga	68160
tgcagaagg	tctgaaaccc	agctcttctc	tcttctgtgt	gatgtcgggg	tgtaatttcc	68220
aggggccagg	agcctgggtc	tgagggcgga	caccaaagtt	ctagtgggtg	ctattagcag	68280
cgtttaaatc	taatggatgg	atgttggtctt	gttaccctgc	tcaaaagctt	tcagcagctc	68340
cccactgtcc	acaggacaaa	aatccagatg	ctagcctggc	attcaaggct	gtcactagt	68400
tgatctcaac	ctctcccctt	ccctctttac	ctcctaccaa	cagcggggca	gagccacccc	68460
ctgtggacca	agattcccag	tctctgggtc	tgtgtgtgca	ccagttctct	tgctgtgggtg	68520
gctcacctcg	cctcagcttg	tgaaatccat	ctggtctgct	gggatcctgc	tcaaaatgtc	68580
atcttctcca	aaaatcatta	ctcaggcttt	ccagcatgtc	tgagtccctg	gcacttggtc	68640
acaccttctc	tggtgactgg	catttgctct	cacatcatga	ccctcccacc	ccttgcttgg	68700
gcagcatact	ccaggaggca	aggtctgttc	tgccttggtc	ctaatttaate	tgtgcttacc	68760
atccacatgg	taccagctaa	ttcttgttga	atgaatgatc	gttgaatgag	tggtattcttg	68820
ttttggcctc	agaaccaatt	agaaggagcc	agaaaaacac	atgggggtgg	gggaggtgca	68880
gtgtggtgca	gtggaaaaaa	acccttctgg	aaatctcagc	tctgtcactt	actttgtcag	68940
ctctgtgact	ttggatggac	cacttctttg	tcagtatggt	gggagaaata	gacatgcctc	69000
tctgggctgt	tgtaaggatt	acaaattagg	tcgagtgcct	ggcatgtggt	gggttgaaca	69060
gatcacagct	agcattacag	atgatataat	aaagccaaaa	aaagatgcct	aatgtccacc	69120
agttggtgaa	cggacaaagg	aaatgtacca	tatttgggat	attattttggc	aatcaaaaaa	69180
agtactgaca	cctgctacaa	cacggatgaa	tcttgaaaac	attagactaa	gtgaaagaag	69240
ccagacacaa	gaaactgcta	atgattccat	ttaaataatga	aatatcgggc	caggggtgag	69300
tggtcatgac	ctgtaatccc	agcacttttg	gatgccaagg	tgggcagatc	acttgaggcc	69360
aggagtctgt	gaccagcctg	gccaacatgg	cgaaaccccc	tctctactaa	aaattagccg	69420
agtgtagtgg	catgcacctg	taatcccagc	tacttggttg	gctgaggcac	aagaattggt	69480
tgagcctggc	aggtggaggt	tgcaagtggc	caagatcgtg	ccactgcact	ccagcctgga	69540
tgacacagtg	aggttccgtc	tcaaaaaaaa	aaaaaaaaaa	ggaaaaagaa	aaaaagaaat	69600
ttccagaata	ggccaatctg	tagaggcaga	aagtagattc	atgattgggt	aggcctgggt	69660
gtggaggcca	tggttagtga	tggctaattg	ggaaggggtt	tcttttgggg	tgatgaaaat	69720
gggtggactt	atggtatgtt	aattatacct	caataaaact	gttattttaa	ggaagaaaag	69780
atgcctggat	tccccaggaa	gtgtacagta	gacttctgtg	agaatcagaa	atgatttctg	69840
gggaagatgg	gcgagaggag	agtaagtggg	agaagtgacc	acgtgcgcaa	ctctcatcgt	69900
tctgccttga	gagccttctc	cctgcaactt	tatttattta	tttattttga	aacaggttct	69960
cactctgtta	ccctggctgg	agtgcagttg	tgtgatctca	gctcactgca	gcctcgacct	70020

gccaggctca	agcaatcctc	ctgtttgagc	tcctgagtag	ctgggactac	aggcgcgatgc	70080
caccacatct	ggctaactct	ttattttatt	atttatttat	ttatagagat	tggggagagt	70140
cactctgttg	ctcaggctgg	tgtcaaacgc	ctggactcaa	gtgatcctcc	caccttggcc	70200
tcccaaagt	ttgggattat	gggtgtgagc	cactgtacct	ggcacctcct	gcaacttctt	70260
cctcaagtgg	aaccaatgag	gaagcaagca	actcagagct	ttcacaagtt	ttgatttcaa	70320
tcagcaacgg	gcttccaatg	caacccttct	ctcctgtaac	cagcctcagt	agagaggaac	70380
tggaggtgaa	ttggccccc	tcacaccccc	acagtgccaa	gctgggccc	tccatcaggg	70440
ggagaacaca	tgccgtgtaa	gggacagcca	acagcataaa	ataggaattg	tgtgatgatc	70500
ccttttaagc	ctattcagcc	cagggaagt	catatgatca	gccccatttc	atagatgaag	70560
aaagtcagg	tcaccatta	gcacattgtg	gggtctggat	ttaaacccag	tctgtctggc	70620
tcccaaggct	acattcattt	agacattacc	ttacttttac	atcttcttct	cttttcttct	70680
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cttcttcttc	ttcttcttct	70740
tcttcttctt	cttcttcttc	ttcttcttct	tccttcttct	ccttcttctc	ctcttcttcc	70800
tcttcttctt	cttcttcttc	tccttcttct	cttcttcttc	ttccttctct	tcttcttctt	70860
cttcttcttc	ttcttctctt	tcttcttctt	tcttcttctt	cttttttttt	tgaggtgggg	70920
tcttcttctt	ttgcccagg	tgaatgcagc	atcatcatac	ctaaatgcag	ccttgaactc	70980
ctggccttaa	gcaatcccc	tgccctgggc	tccaaaagt	ccaagatttc	aggcagagc	71040
caccatgccc	agcctgcatt	tattctcttg	taagaaagat	atcattttaa	acagacgaga	71100
aaataaagag	ggacatgaaa	aagacgcac	accattaatt	ggaccactca	gagataatca	71160
tgggttaacat	gttggtatgt	tccctcccgt	catttgactg	gatgtatgtg	ataattttaa	71220
tgatctcata	agcttttctt	tatgtaatac	aatagtagcc	aaaaacatga	ttttaaatgg	71280
ctgctcacaa	ccccatctcg	tggttctgcc	acgccttggt	tatccccatc	cacccccctac	71340
tccctttccc	cttccctgcc	tgtgtggggg	tcctagatga	cgggtgagca	gagggcagcc	71400
ttggtcagca	gattggagag	tgcaataaat	aaaaacactc	agaaggcgag	ctgtgtgcaa	71460
gtgggcttat	cacaaaagag	caccttggga	tattccagag	aatgacctca	tacccgctaa	71520
tcactatcca	taatctggtg	ctaactgtac	tttagctgaa	ggtgctggca	ggtcctgccc	71580
aggtgctgct	aagaacactt	ctattctgtg	agaatcagag	atgatttcta	gggaaaatgg	71640
gogagagggg	gtaagcagga	gaaacaaccc	acaggcacag	ctctcatctt	tctgcccctga	71700
gagccttcct	cctgccacgt	ggttttgttt	gtttgtttgt	ttgtttgttt	cagatagggg	71760
ctcactctgt	caccaggtct	ggagtgtagt	ggcaagatca	tggctcactg	aagcctcgac	71820
ctcccaggct	caagcagctc	tcccaaat	caaagcttgg	agtgtgtgtg	ccagtgggta	71880
tgtctaggag	ccctttttcc	tgccagcccc	tcaggggatt	gatgactctc	aaatgcttca	71940
ggtgtgacat	gggcacagca	gtgagtcatt	cctctgacat	tctttgggaa	gaacattttc	72000
catccaggct	tccaggcata	agatccagtc	ctctggtgat	aaggagtcca	cagacaggac	72060
aatgtctgag	tgtatcttaa	acccaggacc	atggcttgtg	ttcacaccag	accctccagg	72120
gattttgagg	tgttttgttt	gtttgtttgt	ttgtttgttt	gttttttgag	acagagtctc	72180
tctctgtcgc	caggctggag	tgcaagtggc	cgactcagc	tcactgcaac	cttcgcctcc	72240
cggttcaagc	gattctctct	tctcagcctc	ctgagtgtct	gggactacag	gtgtgaccca	72300
ccacacccgg	ctaatttttg	tatttttaat	agagactgtg	tttcaccatg	ttggacagga	72360
tggtcttgat	ctcttgacct	cgatgactct	ccgcctcggc	ctcccaaat	actgggatta	72420
caggcatgag	ccaccgtggc	ccgcccatt	ttgagttttt	atgttcta	cccaaacatc	72480
tgtctacagg	cccctcagca	tattctttcc	tgggtccagt	gtcacctccc	aggcctgcag	72540
gctggctaga	gcagtagggt	gtgtgggaaa	gctctgggct	ttgcaggcac	tgatcagctg	72600
tgtgacctta	accaccctga	acctcagttt	cctcactgtt	aatggaaata	ggtaaccagc	72660
cagtttgttg	caaggactag	agagtaacct	tgggaataaa	aggtagcagc	agcttgggct	72720
ctggagatgg	actgtccaag	accaacttcc	agttcctccc	cacacaagct	ctggcactta	72780
gattcctggg	acctccgctg	cttcatctgt	aaaatggagt	aacaatagga	atactttata	72840
gagttgtaag	gattgagtgg	ctggatgaac	gtcaagcact	tcaaagggga	cctggcatgt	72900
agtgagtgat	caatataaac	cacctggctt	gtagcaggtg	tgctgtgtgt	ggctgcaggt	72960
gttattagta	acatctgtgt	gcccttcaga	gcgtgcacca	cacttcacac	cttgtggagt	73020
ctggaatgcc	actattatag	ttcaggatag	aaaacctccc	tgcaagcact	cgctttagct	73080
tgtctccacc	gaacaaaaca	acacaagttc	tttattactt	ggaatgggaa	aaactcaaa	73140
gcaaaaaaaa	aaaaagactt	tcgagttacc	ccaaatctta	agccaaagtc	aatgaaaaat	73200
atcaatcttc	atattcaatt	tttgcgatac	ttttgtctcc	ccagcagtea	atggagagaa	73260
tccaagcaca	cagaaatgtc	aattaccagg	ggcagggcta	tgaattcctt	tcagagccct	73320
gggctgggga	agagtgcagg	cagacagatc	tgggtcctgt	tatcacgttc	ttagattggg	73380
tgtccttgta	ggaagtcata	agcatcctag	tgcctttggt	tgctacctat	aatgcctacc	73440
tcagagagta	attaaggata	gtaaggctct	acgtgaaaag	tgctcgcccc	tggcacatag	73500
taggtccttc	attaatggca	gctactaatt	tttattacat	acgcaaaatc	acattacagg	73560
tcaagtacgc	tacatgacag	tgaacaggtt	ttttgttttg	tttgttttga	gacagagtct	73620
cgctctgtca	cccaggctgg	agtgcagtg	cacgatcttg	gctcaccgca	acttctgcct	73680
tcaagcaatt	ctcctgtctc	agcctcccca	gtagctggga	ttacaggcat	gtgccaccac	73740
gccagcta	tttttttggt	attttttagta	gagacggggg	ttcaccatat	tggccagact	73800
ggtctcaaac	tcctgacctt	gtgatctgcc	caactcagac	tcccaaagt	ctgggattac	73860

tggcatgagc	caccgcacct	ggctgtgaaa	cagtttttatt	gtgttttctgt	ggaatgtgtc	73920
ctacccaacc	tatagctaac	tcctatagtt	ccctcagttc	tcagctcaga	tatcccttcc	73980
ttctctgact	gttacctagt	actggttttc	atagcaccag	gtacctctct	ggcatagagc	74040
ttgtcacagt	tgcagtttaa	tgtaccatca	taggatttta	aaaatattca	gttgtgtctt	74100
ccattaggct	ttcatttggg	aactccacgc	aggcagcagc	tgtatatatt	gtattgccta	74160
ctgtatcctg	agaactttgt	accctactta	gcacagaatg	gaggctcagt	aaatactgga	74220
catgagagag	agagagagag	agagaggaga	gggagagaga	gagagagaga	ttcaacctac	74280
aatcccagct	ctgagcttct	agttccctga	tgggtaggac	tgtgatgtgt	ctcacacggt	74340
aatgagcact	tatgcagaag	aggctcagaa	aatttctcct	catggccaac	ggaagactta	74400
gagttctttt	ccaagctcca	ccgtttgtct	gcattgcaaaa	tttggactat	cacttaagtt	74460
ttccaagcct	tgctttttct	atccctaaca	taggacaata	ttcagcattg	gtgtttgttt	74520
gttgggggca	ccatgtttca	ggcacttagt	agattattgt	accaccacat	ttcaattggg	74580
cctcctcaag	ccctgcaaca	tctgtgaggt	gggtatcctt	aacaactcac	agatgagcaa	74640
caggagactg	gggggatgag	ggaactgcc	aggaggtcca	gcttatgggc	agcagagcca	74700
agaatggaac	caggggtctt	tattttttta	tttttttatt	tttatttttt	aaccagggtc	74760
ttttaacatc	cgaggaccac	attcctttgt	ctttccaaat	catcacctgc	cccatgcaac	74820
ttacagggtg	agttacatta	aacaacgtat	gtaaatggct	ttgtgctagt	tattccacac	74880
cacaggggaa	gtgagtcacg	gacaagagt	gcacgcgtcc	attcggatcc	tggctctgac	74940
acttacctgg	aaaatgactt	aaccattccc	aggatcagct	gtttgtctgt	aatttaggta	75000
gtttaatggc	acttgtgtcc	tagagttggt	tagaagggtg	aataatatgg	agcacttaac	75060
atacttagca	cctagaaaca	cttcctaaat	attagttgct	gctgttggtt	tcgttattaa	75120
aattttctgcc	taagatctca	tttcaggagg	cccaactcaa	tctttgacaa	gcttaaacaa	75180
aaattgcttt	tcttcattta	ttcacttaca	cagcaaacat	gaattgagcc	tgtactgtgt	75240
ttccagaact	gtgcaggacc	agagaggcac	aggtgaagga	agcaaggctc	tggctctact	75300
ggggaaacag	caagaagatt	gctacaatga	gtgtgggaaga	gggctggact	agagagaagc	75360
cctgattagt	gtccttgcta	cctttctctg	ggagagccaa	ggcaggcttc	ctggaagagg	75420
tgatccttgg	ctgaaacttc	gatgaagaaa	aggaaagagc	gcagtgggta	gggaggaaag	75480
ggcattctgg	gcagatgaaa	tgacatgtga	caaaatatgg	gtgatcannn	nnnnnnnnnn	75540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	75600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnngca	ctccagcctg	ggtcactgag	tgagagaccc	75660
tgtctcaaaa	aaattaaaaa	aaaaaagtc	agaagaacat	ttgggtctca	ctctgtggcc	75720
caggctggag	tatagtggca	caatcatagc	tactgtcaca	ttcaaactcc	tggcctcaag	75780
tgatcctcct	gccttagcct	tgaataaagc	ttggattaca	gatgagccac	cacacccagc	75840
cagaccatta	ttcataatag	ccaaaatgtg	aaaacaaccc	aaatctccat	caactgacaa	75900
atggataaat	agaatgggtg	ttgatccata	caatggagta	ttactcagc	aataaaaaag	75960
agtccctgata	catgctacaa	ggatgaacct	cgaaaacatt	atgctaagt	aaagcagcca	76020
atcacaaaa	gctacatatt	acaagattcc	atttaaatga	aatgttcaga	ataggtaaat	76080
ctaactttta	tcacaggcaa	agctatgaca	ggaaatagat	gagtgggttc	ctagtgcctg	76140
ggggcgagagg	tgggggtgag	gcgagtgagt	actgctaatt	gtacagagtt	acttttgggg	76200
ataaagaaac	tgttctgaaa	tggactctgg	tgatgggttc	actactctga	acatactaaa	76260
actgttaaat	tatatacttg	aaatgggtga	cttgtgaggc	atggaaatta	tatcttaata	76320
aagctgtttt	acatatttta	catattttaa	aatgcagggt	gagggatgag	ccctctaaag	76380
agaagcagga	gtttgaggag	gttctaataa	ttgtgtgggt	ggtactgagg	catataaaat	76440
tgtcagacct	catcaaaatg	tatgatgtaa	tcttcaagaa	agttgatttt	aaagaaacac	76500
caccagaccc	aggtggagaa	ggcaggaaga	agttacacaa	ggggtaggct	aagagtgggt	76560
gctcatgtct	ataatcccag	cactgcggga	ggccgagctg	ggtgggtgac	ttgaggtcag	76620
gtgttcgaga	ccagcctggc	caacatgggt	aaagcccgtc	tctactaaaa	atacaaaaat	76680
tagccagacg	tgctcgcgtg	aaccagggg	gagaagggtg	cagtgaagcg	agatcatgcc	76740
aatgcactcc	agcctgggtg	acagagtgg	actctgtctc	aaaaaaaaaa	aaaaaagtta	76800
catagggggg	agtggcagg	gtcaagggca	ggcagggtct	ctcctatctc	caggataaac	76860
tcatagggga	cttagatgcc	atgtgggtcc	ctaatagccc	tccacttggt	tcttgagacc	76920
actcttatgt	gtatcatttc	atgtcaggcc	tcttcttccc	aaccacccca	gccatcccag	76980
cctggctgcc	aaccccacct	cctccagccc	ctgtcacccc	ataattgggg	ccaggaggca	77040
tgggagagtc	gccatctctc	ggtgccatct	gttgcatctt	tacagataac	catggctgga	77100
tgcggcagat	cctgggggtg	agcagccgct	gttcagagca	gtgatcaaga	cctccccatc	77160
tccacccctc	aaggaatcgg	ttttcttcca	tagccacatc	aggtgctgtg	caggaaggag	77220
ttgaaacgag	aagccaggag	caacgagaag	gacactaaca	tttattaagc	ctgcagact	77280
ctcacagcac	tcccacggaa	tcgatattat	tatccccatt	ctgaagacca	ggcaactgaa	77340
gctcaatgtt	taaggaaact	accgaagtca	ccaactgata	aaagtgatgg	aagctgggat	77400
tcaaatacaa	gctaaacttc	cttccaagct	tactccacaa	cacagaggtt	ggggaaaggg	77460
gataaaaaag	gaggggagcc	caattccatt	tccacccagc	tcctgaggcg	gagcttgtca	77520
gcacagctct	ctccttccca	gaataggaag	atacccatca	gaggcaagtc	ctagacacca	77580
gcagtggtaa	ctccctgccc	caaggcagct	gcagacagcc	tatggctgta	gttactgctc	77640
ccaaagagt	ttagaattcc	cactcccagc	ttcggggcca	ctcacacaag	gtgattgaa	77700

tggaaccag	agactctcca	caatgcctc	ctagagtaaa	tgaggctatg	taactttgtc	77760
caaatgagta	atttgaaaac	ctgggggctc	ccagctcctg	aaaagggag	gatgtggggc	77820
cctttatatt	catactccac	tttgtgcagc	tctcccttgt	cttatgatag	ccctattaag	77880
aaattcctct	cccagcacgt	ctccttcaaa	gagctctaga	cctgaggctg	tcagaggctt	77940
aggactctgc	ctattagtcc	cagggctctg	atgaccagca	ggacacctgg	cattcagtga	78000
ccactggatt	agataaatga	aacagtgggc	agagtgccac	ccaatctccc	cctgaagttt	78060
gaagaggctg	agaagtgagg	ctgtccaact	gctgacctg	ctttctgtcc	acctggccac	78120
ctaacccttt	ctggcttcca	cctgcccctt	tgccatccct	ccccccagcc	caccagccc	78180
attttcaggc	atacctgggc	acgtgctgga	atagaagccc	tcgttcttca	gaatgatcaa	78240
cagggagccc	cagcccagga	gtacagcaga	gaagaagagg	ttctccagca	cagccgtgca	78300
ggccatccac	cagcgctcc	ggtacgcctg	ttgcaagctg	ggggccatgc	tggccccgag	78360
cctgcacaga	aacagagcgc	tgggtgaagg	gccccccagt	ggccccaggg	aagggctctg	78420
catcatgggtg	gcacccgaga	cctctcgggc	cagcccgcga	ggagcccctc	atggaggccc	78480
catagagccc	tgggcttccc	agccggtgcc	aaggagctgg	ctccgcgcgc	actagcagt	78540
ccagaggctg	acgcggcacg	gggctcccgc	tgagccacta	tcggaaacaa	ggaaggctct	78600
gtctgcgcgc	tgcagcttcc	tagcaggctg	ccgggttctc	tcacccaggg	cagggcgctc	78660
agggccgggg	tgctggggag	aaagtccgca	tctgcccagg	tcccagagg	acagcaagg	78720
gcagagcgcg	ctctgaagca	cgcggggccc	atgtccggac	tctcgcgcca	ggaaagaccc	78780
ctagaagctg	gcaggaagaa	gggcaagttc	aaggctaccc	tacgacccca	tcttccagtt	78840
gcccctccaa	gacctctcct	tccctctggg	gccgggagac	agcaagccct	ccccctttcc	78900
gtatcagggtg	acccacgacc	ctacagtctc	tccggccaag	ccaacagctg	ccacgtggag	78960
ggagacccag	gacgggctct	cctcggttcc	ctcctcccc	gcgcgcccct	cactcactcc	79020
gcagggctcg	gggcaccagg	ctttgcacct	cggaaacccg	ttgccccct	ccagccccgg	79080
gagggggctc	ggacttcggc	aggaagtctg	cgcggtctg	actttataag	ggcagcggtg	79140
gcggtggggc	tggcgggcgg	gtgtgtttac	caaagggagg	gaaagagccc	cagctcccc	79200
cgccgcggcc	gctgcagcct	cggcgggagg	agagggaacg	cgggcagcgc	gggggcgggg	79260
agcgacaact	gggatgagac	cgaggaaagc	ggagaggaga	agggcaagaa	agaccagag	79320
agaggggagg	aagtaccagt	cacttcttcc	agggggactc	ggtattctca	tctgtgaaac	79380
ggggctttgg	gttcaagcgc	tccaggaggt	ccgctggaac	tctggcaaac	gcgcagctct	79440
aagcagagga	agtgcagcga	gcggggaccc	gggaggaaga	gaagagtccg	aggggtcaga	79500
gaaaagaaaa	gggaaggacg	cgcttggcga	gatgggacac	tgtgcgcggg	gaccgcgggc	79560
gcaagtaacg	gtctttcctt	gggaagcctg	gcagtgtcgg	cgggagccgg	cctcggtgtc	79620
tctcagccga	cgcatagccg	gagaccctac	gcgcgcccc	tccccgccc	cgctgctcac	79680
ctccggtcac	cggcaaatga	gcagccagca	gctgcggacg	cctccgggag	cgcaacgctt	79740
tcgcggcgcg	tccggagtc	cgtgggcccc	gccctgagcc	gcgcggcgcg	tgggggtctc	79800
tctgcgtgca	ggacccggcc	gccacggagc	ttcagcctga	cagcccggtg	gcctcgccctc	79860
cgctgtctcc	tcggaagaag	cgggggaact	gggaaccgcg	cgggcgccag	aggtctcgca	79920
agctgggctt	ggatgaagtg	gatctgcgga	gttgatagtt	gtatttacac	gcgtccggag	79980
ctgcgccccg	aggtgggggc	gggggctccc	ttcttttccc	ctccccctag	gtcgagtttc	80040
acgcgcacgt	gactcgcccg	ctggctcccg	acactctccc	tctggcacag	ccccagcacc	80100
tacatttcca	ccctggaccc	ccatcttctc	ccccaaagcc	ccagactaac	atcaggcagc	80160
gccctctgta	tccttggtca	aaacaaagt	cgattcggct	gaagccgact	gaccgcgatt	80220
cagggccgcc	ttgggtgggg	ttttgaactg	tcgagctgga	agcagtgttt	tccgagaggc	80280
agagtggcac	gggtttcttt	ggagttagtc	agactcgaggt	ctgagctctg	actttttaac	80340
tgactacctc	gggttaccta	gggcaagtta	cctctctgag	cctcagcttc	ctcctcttta	80400
aattcggtta	aaatggaacc	tacctaaact	cccaaaggaa	tcgcgattgt	gatgcaggta	80460
aaatgctaag	catagcattt	ggcatagtaa	gcataatgtt	aattgttgct	gctgtcatta	80520
tttcagaaga	cctggtgatc	ggatgcttcc	agatcaacaa	ttgattgact	ccaggtaaat	80580
ctctcagcct	ccctgagcct	cagtatcctc	atctgtaaaa	tagactacta	tggtgtggag	80640
taatgagaag	taatctcatt	acatgtgagt	ttaattgtgt	gttaagagt	ctgtcaatgc	80700
atgctgagct	taatacctag	gtgatgggtt	gataagggtc	ataaaccacc	atggcataca	80760
tttacctacg	taacaaacct	gcacattctg	cacatgtacc	ccagaactta	aaataaaaaat	80820
aaaatttttt	taaaaaaaga	gtgatactgg	tggccagggt	tggtggttca	tgctgtaat	80880
cccagaactt	tgggaggcca	aggcaggagg	atcgcttgag	ctcaggagtt	cgagaccaac	80940
ctggacaaca	tggtgaaacc	ccgtctctac	aaaaaagaaa	aaaaaatagc	caggcatggt	81000
ggtgtgcacc	tgcatgtctc	gctacccagc	aggctgaagt	gggaggatca	ctgagctgga	81060
gagatggagg	ctgcagttag	ccaagatcat	gccactacac	tccagcctgg	gtgacagagt	81120
aagactctgt	ctcaaaaaa	aaacaagaat	gactacagaa	agctccaaga	aggcctcaga	81180
taaaagggaa	cccctgaaca	gatgagccac	caagccaaga	gaggaactaa	tggctaccat	81240
agacagggca	ctttccaaaa	taaaaatact	gttattaatt	cctcaagaca	tcatgggtccc	81300
atttaaaccct	catagctttt	cacagaggga	gaaactgcag	gcttgaagct	ggagcaagggt	81360
tagaggtagg	atgcagagtc	aggtcggcct	ggcatttaag	tacggctcct	tccattcctc	81420
ccagaaggag	aatggcaaga	gcaaaggcct	agctgtggga	atggcacaag	gagttctcgg	81480
tggccaaagc	acatgtcagg	ctctgatgggt	ttaacttctt	aaaatgcaat	actgcctccc	81540

agaacttcca	gatcaaggctc	aaactcctca	gctctacaca	gggggaccta	gagtcaactt	81600
tctaagctag	gagagtcag	gatccctttg	agaatacaaa	agacagtggg	cgcggtggca	81660
gtggctcatg	cctgtaatcc	caacattttg	ggaggctgag	gcaggaggat	cacttgagcc	81720
caggagtcca	agacctgctt	gggtcaacata	gtgagacccc	tatttctaca	aaaaattcag	81780
ctgagcatgg	tggcatgtgc	ctgtagtctc	agttactggg	gaggctgaag	taggatgatc	81840
cctgagcctg	ggaggtccag	gaagctggag	tgagccgaca	tctcgccact	gcactccagc	81900
ctgggtgaca	gagaccctgt	ctcaaaaaaa	aaaaaaaaaa	gaagaaatat	gttattgatc	81960
tactcttgac	aaaaatgctt	gtgtgaacat	ggacacacac	actcatcaac	attcacattt	82020
caaggttttc	atggaccctt	tccatgaggc	tctagtggtc	catggacccc	catggctgga	82080
acacttgctc	ttcctcatct	caaccacat	ttccatggag	ttggactgtc	tgctgcatga	82140
ggacacaggg	ctcatttggt	gtgttcattc	actgctgtgt	atcccagcac	ccagaacagc	82200
acctcaccta	aggggcactc	agcacatgtg	cagtgaagag	tcagtcagct	ggtttcacac	82260
ctcccagctc	ttgcacctgc	tattccttct	tgtgggaatg	acagatttcc	ttcattttct	82320
tttttttttt	ttttgacaga	ttccagctct	gttgcccag	ttggagtaca	gtggcacgat	82380
ctcagctcac	tgcaacctct	gcctcccagg	ttcaagcaat	tctcatgcct	cagcctccca	82440
agtagctggg	attacaggtg	cacaccacca	cctgtgagct	gataattttt	tcttttcttt	82500
tcttttttcc	tgagacagag	tctcactctg	ttgccaggc	tgaggtgcag	tggcgtgatc	82560
tcggctcact	gcaagctcca	cctcccgggt	tcaagtgatt	ctcctgcctc	agcctcccaa	82620
gtagctgaga	ctacaggcgc	gcaccacat	gcctggctaa	tttttgtatt	ttttagtaga	82680
ggcgggggtt	caccatattg	gacaggctgg	tctcgaaact	ctgacctcgt	gatccgccca	82740
cgttggcctc	ccaaggtgct	gagattacag	gtgtgagcca	ctgcaactcg	ccattttttg	82800
tattttttta	gtagagatgg	ggtttcacca	tgttggccag	gctggtcctg	aactccttggc	82860
ctcacgtgat	ccaccacact	tggccacca	aagtgttggg	attacaggca	tgaaccactg	82920
cgctcagcct	ccttcttcat	ttctaattga	ctcatccttc	acaactcagc	tcaagtttca	82980
cttctctctg	gaagctctac	tctaggctgg	attcaggggc	ttgtccacat	acccacccaa	83040
tactctgctt	acctctatgg	aagtccccac	actgatctag	aataatcagc	ttagttttct	83100
gcccccatcc	cgccccatga	gatgtacatc	ttgtgggggc	aggaaccacc	acgtggtagg	83160
tgatttgtgt	gcctgctgcc	tatcacaggg	cctggcgctt	aataagcttg	cggccaaact	83220
ttgttgaata	aatgaaaagg	gaatggtggg	aaaggaagct	gaaaaggtag	gctaaaatca	83280
gttttgaatt	acctctggga	ggccaaggac	tttcagctct	gcagggttag	taacaggaaa	83340
ctcctggatt	ttgttttctt	ttggttttgt	ttgtttttaa	tgaagggtag	cgttatcgtc	83400
aggtttttgt	gtttaattaa	tggagcatat	attggaaaagg	acagagacct	taaagcagtt	83460
aggagaccac	cataatagtt	cacatttttg	agccataaaa	aggaatgagg	ccaggcatgg	83520
tggctcactc	ctgtaatctt	atcacttcgg	gaggttgagg	caggcggtac	acctgaggtc	83580
aggagtttga	gaccagcctc	accaacatgg	agaaaccccc	tctctactaa	aaatacaaaa	83640
ttatccaggc	gtggtggtac	atgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	83700
aatagcttga	atctgggagg	cagaggttgc	ggtgagccga	gatcgtgcca	ttgcattgca	83760
ggtagatgga	tgaagctgga	agccatcact	ctcagcaaac	taacacagga	acagaaaacc	83820
aaacaccgca	tgttctcact	cataagtagg	agctgaacat	tgaaaacaca	tggacacaga	83880
ggggaacatc	acacactagg	gcccgttggg	gagtgggggt	tggggggtaa	ggggagggaa	83940
cttagaggac	gggacaatag	gtgcagcaaa	ccaccatgac	acacgtatac	atatgtgaca	84000
aacctgcaca	ttctgcacat	ggatcctggt	ttgtttttaa	aagaaataaa	gaaaaaacca	84060
agaagaaaca	aacaaacaaa	aataattccc	atttaaaaca	ataaaaaata	ggccaggcat	84120
ggtgactcag	gtctataatc	ccaacacttt	ggggagccaa	cgcgggcaga	tctcttgagc	84180
ccaggagttc	aaggccagcc	tgggcaacat	ggcaaaaccc	tgtctctaca	aaaaatataa	84240
aacaaacaaa	caaaatagcc	aggagtgggt	gtgcatgcct	gtcatcccag	ctactcaggt	84300
ggctgagggtg	ggagaatcac	ttaagcctgg	gaggcggagg	tagcagttag	ctgagatcgt	84360
gccactgcac	tccacctgga	gcaacagagc	aagattttgt	ctctaaataa	ataaaataaa	84420
taataaaaaa	cagagaagag	gaaagacacc	tgagatatat	ttccatatct	gaatcaatag	84480
gattttatcaa	cgttctcctc	tacccccaaa	actaatcctt	tcttaaaact	tgttctcctg	84540
acactactca	taggttaagt	ataacagcat	tatcacattg	gctgtcatgt	gggctcctgg	84600
ctagaggctg	cttcacagct	taatggacaa	gagcactgag	acagggtggg	tctaaatcct	84660
ggctctgcag	ctgattattt	gtgtgatttt	gtccaaatca	ctccatctca	tgagcctcac	84720
tcttctagtc	tgttaagtgc	tgaaaataaa	agtatccaat	tcaattcatt	atttaatgaa	84780
ttattttagcc	taacaaatag	ctattataaa	tattttaggt	gggcacagtg	gctcacgcct	84840
gtaatcccag	cactttggga	ggccaagggt	ggcagatcac	ctgagtcagg	agtttgagac	84900
cagcctgacc	aacatgggtg	aaccccgctc	ctactaaaaa	tacaaaaaatt	agctgggtgt	84960
gggtggcatgt	gcctgtaatc	ccagctactc	aggaggctga	ggcaggagaa	cgcttgaacc	85020
caggagacag	aggctgcagt	gagccaagat	cgtgccactg	cactctagcc	tgagcaacag	85080
agcaagactc	tgtctcaaaa	aaaaaaaaaa	aatctctgca	tgaagaatgt	acataaaatg	85140
gtgcagccat	ttcggaaaaa	agtttggcag	gtcctcaaat	agttaaacat	agagttacca	85200
ctatagccca	gcaattccac	tcctaaatat	actacaccca	agagaattga	gaatatttgt	85260
taacacaaaa	atgtgtatac	aagtatttat	agctgtatta	ttcattacag	ctaaaaagtg	85320
caaacatccc	agcagtcctt	cagctgatga	acggagaaac	aaaatgtggt	ataccatac	85380

aatgtcatat	tatttgGCCA	taaaaaggaa	gtactgatac	atgctacaac	atggatgaac	85440
cttgataatg	ttattctaag	tgaaagaaac	cagacacaaa	agaccacata	ttgtatgact	85500
gcatttatat	gaagtGCCCA	gaataggcaa	atccacagag	acagaaagta	gattagtggg	85560
tgccagagac	tgaggggagg	agataatggg	aaatgtggaa	tgactgctaa	tggtatgggg	85620
tttcttcttg	gggtaatgaa	aatgtttgtac	aattagataa	tggtgatcat	tgtaaaactt	85680
tgtgaatata	caacatgctg	aattttatac	tttattatat	tttatttttt	ttgagacaag	85740
gtctcgctct	gtcaccagcag	ctggagtgcA	gtggcacgat	ctcagctcac	tgcaatctct	85800
ctgcctccca	ggctcaagca	atcctcctgc	ctcagcctcc	tgagtacctg	acactacagc	85860
atgtgctacc	atgcctggac	aatttttgca	tttttagtag	agacagggtt	tcgctatgtt	85920
gcccaggctg	attttgaaact	cctggactca	agtgtctccg	ccacctcagc	ctcccaaagt	85980
gctaggatta	caggtgtaag	ccaccactcc	cggcctaaat	tgtattcttt	aaaagactga	86040
attgtatggg	gtgcgaatta	tatctcaatt	taaaaaaaac	aaaacaaaac	aaaaaaaaaa	86100
cctttgctg	tgtcaggcac	tagggattcg	atgctgaata	agacacagac	cctaccctca	86160
gagaacacag	agcccagcag	gagagagtca	cagatgaatc	aagtgttaca	tcacttatag	86220
gaagcgccat	ggaagaaaga	catggtgcca	tgagaacata	cgcttagaga	agggaatttc	86280
atctagactg	gggctcaggg	aggaatcttt	cagggtgatg	cttgtgctca	gagttttcca	86340
tgtcaagaatc	agtagaattt	atcaatcctc	cagaggagga	aacagcaaat	gaaaaatctt	86400
acaacaggag	gatgCGGAGA	cattccgaga	gctgatcaag	ggctgggtgtg	aacaaagcac	86460
ataggatgca	gagcctgtgg	tgtgaggttg	cagctggaaa	ggtaaaacac	taattacatt	86520
ggatcttctg	agacaataaa	gagtatgcaa	taatctcaaa	cgaccgaaac	tgaccttctt	86580
cctccctaac	ttgcttgctt	ccactgttgc	ccgtatcata	aaagcaccac	cctcttctac	86640
ccagtggctt	aagacacgaa	actcaagtca	tcccaggctt	tctcccacc	tcactctcca	86700
catccagcct	atcagcgagc	ttgtgggtct	taccacgtaa	agacttctca	tctccagcta	86760
ctaccatccc	ccaagcccag	atcaccatca	gctcaggcct	ggactcctgc	aacctttcta	86820
accgggtctt	cccaatccta	cccccgcaac	atgaccccaa	tagcccatca	gaatggacta	86880
atcgagatgt	agatttgatc	aggccacatc	ccttgaaagg	cttctgtgta	ccctcgggga	86940
aatgcacaaa	ctcccaatga	tggcccttga	gtcctgtgcc	atctgggtct	gccctctgcc	87000
ctctgtgtct	ttgccatggg	aacctccttc	acaccattta	atactccatg	ctctctccta	87060
cctcaagttc	ttcctgggct	ggaacattct	ctgcactagc	ctagccaact	aaccttttag	87120
atcttttgtt	ttgtttgttg	ttgtttgttt	tgtttttag	acagtcttgc	tctgttgcca	87180
ggctggagtg	caatgggtgca	atctatctcg	gctcactgca	acctctgcct	gccgggttca	87240
agcaattctt	ctgccttagc	gtcctgagta	gctgagacta	taggcaccta	ccatcacgcc	87300
cggctaattt	ttgtattttc	agtggaggtg	ggttttcacc	atgttggcca	ggctggtctc	87360
gaactcctgg	cctcaaatga	ccacctcctt	cggcctccta	aagtgtctgg	attacaagca	87420
tgagccactg	tgcccaggca	acacttcaga	tcttaatgat	catttctctt	aagtgcctga	87480
cctctttagt	taactagcct	gactccagca	atgaatcctt	ttgcaatgta	acctatatata	87540
catctgagtt	tccctttgat	aaaactcatc	atatatttgt	tctcttgaca	gttcagaggg	87600
caagggcctt	tgcccacctt	cctcaccact	atcctctcac	cacttaacac	agaactcacc	87660
accaccatg	cctcctgcct	gacaaattcc	taaccatcct	tcaaactctca	ctcacctatt	87720
accttctggg	aggcagtctt	ccctgagcac	caagacaatg	ggacacattc	ctttatacac	87780
cctgctgaac	atctcttttt	tgaggggCGG	gtagagatga	gtgtctcact	atgctgccca	87840
ggctgacctc	aaactcctgg	cctcaagcga	tctcctgccc	ttggcctccc	aaaatgctgg	87900
gattacaggc	atgagccact	gtacctgacc	gcaactgggt	tagnnnnnnnn	nnnnnnnnnn	87960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88020
nnnnnnnnnn	nnnnnnnnnn	nnncctgggc	aacaaagacc	cttctcttac	aaaaaaaaaa	88080
aaaaaaaaaa	aaaaaaaaaa	aaattatttt	aaattagcca	gacatggtag	tgcatgcctg	88140
tagtccaagc	tacttgggag	gctgaaggga	gaggatcact	tgagcccagg	aggttgaggc	88200
tcagtgagc	cgtgatcgta	ccactgtact	ccagcttggg	caacagagtg	agacctcatc	88260
cctaaaaata	aagaagaaaa	tatggcaatt	tgactgtaca	tctctaattg	gatatacctt	88320
aaggatgaga	aaggaaataag	gaaggacaga	aaaaaggaaa	caaagaagta	gcaacagtat	88380
ttagcaattg	tattgttatc	aagtaacatc	aatattggta	aaaccagtaa	ttatatttta	88440
aatactatat	atgtgtatgt	acattttacat	atgcataatg	taggaaccaa	gtttatcaga	88500
ggaagagaaa	gggctacaaa	tgtaaaatca	aggaaataaa	aatttgaata	aaaatatcag	88560
tattaagtat	ttatgatatt	tttcttataa	aaaaattata	tatatgttaa	ctctatccaa	88620
aacccaaaaa	cagtgacaac	ccaggagcaa	taaaaaacct	cagcatccag	actgtagtct	88680
ctaccatttc	caattaaaga	aacccagggc	tagttgggaa	aaatgacaat	ttcatgtcta	88740
gggcaagaaa	cacacctagt	gaaatggacc	tgaacattta	attgtgttag	aaagtaagga	88800
aactctctag	aaataatgtg	atttcatcta	aaagacacag	attctgggct	ggtaaagtgt	88860
tcaatggcca	aaggtgagac	aatttgagca	tcaagaagaa	tcatgacaga	acagattaaa	88920
acatgtcaaa	tatattttaa	aatgaaatat	tataaaagaa	acaattagta	gccatccctg	88980
aaggtcacta	gggcaccaac	tcatatttca	aactggtaaa	taaagtgtga	agccaagcat	89040
ttatttctgg	gtaacaaaaa	agtaaggaat	gtttttcttt	ctagaagaat	tctagtgtat	89100
aaaagtagaa	gatagaaata	gaaaatcatc	cttttgGCCA	ggggcagtgg	ctgtaatccc	89160
agaactttgg	gaggccaagg	caggtggatc	actggagatc	aggagtgtga	gaccagcctg	89220

gccaacatgg	tgaaccccca	tctctactaa	aaatacaaaa	attagccagg	tagtgggtgc	89280
ctgtaatgcc	agctactcgg	gatgctgagg	caggagaatc	gcttgaacct	gggaggcgga	89340
ggtcgcagta	ggctgagatt	acgccactgc	actccagcct	aggcaacaca	gtgagactct	89400
gtctcaaaaa	aaaaaaaaaga	aagaaaagaa	aatcatcctt	ttgcgatcct	aatgaaataa	89460
tgggcctagg	cattgatcat	taatggctcc	taaaatcact	aagtatatgg	ttgatgggaa	89520
actttatagt	ggatggatca	gactcgcaat	gtctaaacca	gttgatcaat	cttaacatcg	89580
taacaagaca	acagacacca	ggggctgctg	acaggagaac	agaggaaacc	catagctcta	89640
ccactgagtt	attcacggca	aaaaaaaaaa	aaaaaaaaatt	aaactgcgtt	tcctccaagc	89700
ttctaactct	gttgtttaca	ggaaatcccc	aaggaaagga	atacttttaa	atgacacatt	89760
aaaacaacgc	caaatccaaa	atatggggaa	atgacccagt	ttcttcaaca	aataaacaag	89820
aaaaggtagg	ggggaggact	gttctagatt	ttaaaagcta	tagaagacac	agcaacccaa	89880
tacactgcat	ggaccaggca	tggatcctaa	ttggaacaaa	ccaactgtaa	aaggatgtat	89940
ttgaaatgat	tggaggaatt	tgaacagtga	ctgcacagta	gatgatatga	agaaattatt	90000
gttatttttt	aggtgtgatc	atgattttat	gggtgatgtt	aagtaaaaga	ggccttattt	90060
gttagagata	catgtacggg	tatagagaaa	tatttacgga	tgaaatgata	cgatgtctga	90120
gatttgcctt	gaaaactcta	gcaggtgtgg	gagaagcagg	tgcatgggtg	ggggaaggga	90180
tagatgaaat	aagtatgcaa	aatgttagtc	tacttttgtc	cctcctgacc	cagcagggtta	90240
aaatacctca	gcatacctct	actcctccaa	ccaggtccaa	ggatcaggcc	aaaactccct	90300
gatgtggtaa	acagcctgac	cccttcttac	ctctctctct	ccagccactt	ccctaagatt	90360
ccccagtgct	ctgtgcccta	gccagcccca	ctcatctgcc	cagattcctc	aatgtttcac	90420
tctctcattc	accattttga	ccccactgtg	ctcctggggc	actctccaag	gccccgcctc	90480
ttcatctcct	ccctccttac	tcatecttca	ggctctggct	taggtgccat	tcctctcagg	90540
aagccttccc	tgaccaccaat	cccataccta	cctagaacag	attatgtgcg	ctctcttggtg	90600
ccccccatgg	ccccctgtgg	gtttgcttca	cggattataa	ctgcctgact	acctgccttt	90660
ctccaccctc	tagactgaga	acaccttgag	aaaaagaaca	catctatctt	gtctgtcatt	90720
gaatccctgg	tgtctggcac	catggctgac	acataactat	cactcagtga	ctagtgtttt	90780
aatgaatgaa	tgagtgcac	tagacagggt	taagaacaaa	agagaagacc	aggcatgggtg	90840
cttcacgcct	gttatcccag	cattttggga	ggctgaggcg	ggcagatcac	ccgaggctcag	90900
gagttcaaga	ccagcctgac	atgggtgaaac	cccgtctcta	ctaaaaagac	aaaaattagt	90960
gggggatggg	ggcacacgcc	tgtaatccca	gctacttggg	aggctgaggc	aggagaatct	91020
cttgaaccca	ggagggtgcag	gctgcaatga	tctgagatca	caccactgca	ctccagcctg	91080
ggcaacagag	taagactcta	tcacaaaaaa	aaaaaaaaaa	aaaaaaaaaa	gagcgagaga	91140
agatgtcatg	gggtaaatga	agacctccct	tcctgggtcc	ctgaccagcc	cctgcccctcc	91200
cccgcattct	acctgtcttt	cttgcttccct	tctgggtact	ctgtttaagc	cgggtccatga	91260
gcaggccatt	ccagggggca	cacagcactc	cgaactgagt	gaaggcaag	gcatttgtgt	91320
aggtgctgac	tgcagaagga	agagagaggt	tgggtgatga	gaagtttcca	aaactccctt	91380
ccaggcaggg	actctcccac	cttacccttg	tctgcatgtc	cctcctcccc	acaccatcag	91440
acctcctct	ggtgtgtaca	gccctgctgg	gaggctctgt	gttcccagct	gggacatgca	91500
gatgggctac	ctcccagccc	taccacatac	ctcgtgccat	gtccccaccg	gccatgttgg	91560
tcagcaagga	gttgagagt	ccaatgaaga	ggtagtgcca	caactgtatc	acagacagcc	91620
acaccagggtg	ccaggcaaag	cgccgagaga	aagcgtagct	ccagaaggag	cggagttcct	91680
gcttctgccc	tgcccctggg	gtctctaata	gggagaggag	gatctgggag	tgaattacga	91740
ggaaagtggg	caggtaggat	ggggagtgtg	gaggcttcaa	tggaacattt	cagatccggg	91800
cccaccttct	accttggct	cccagaactc	accggcgctg	ctctgcacct	gcctcggact	91860
cacacatccc	cacctccctg	ctttgtcatg	ctggccctac	caccttggat	gacctctgt	91920
gttcttctct	attgaaatcc	gatctgtctc	tcacagcctg	gtcaatgaca	cttcttgcag	91980
taataccttc	ctgatctttc	tcagcgagaa	aggtgaaagg	aacgacaagg	agaggagaaa	92040
gtcagaaggg	agaggagaat	gagtgtggat	actctgttct	aacctgctcc	tcagcacctc	92100
cctttctttt	gataecagta	tcctgagttt	ctttgggaaa	tcttctctta	ccctaattct	92160
catggtccag	atgggaccat	gaattcagtg	ttctgttctt	ctctcaaggt	taaccaatga	92220
gatggttctt	cctaacaagg	cagaccggcc	atgagtttag	attaggatgg	acttaatcta	92280
aaataggtcc	tacaccctgg	caagttcaat	gtcctccctt	gattttggaa	gcttcccaga	92340
accttattct	tctttcttta	aaataaataa	ataaatacat	gttttggatc	caattgtcag	92400
atggtaaaaa	taaaaacaaa	aaaatcaatt	ttattctgta	tatttaagat	atacaatatg	92460
aggtcatagg	atacatatag	ctactaagat	ggttactaca	gttaagcaaa	ttaacatatc	92520
catcatctca	cttttctacc	tgttttgtga	caagagcagt	taaaaatctac	ttgttttagga	92580
aagtcccaaa	cacaatgcag	tttgatgacc	tacagctctc	gtgctgtgca	ttagatctct	92640
aggcctgttc	atcctgctca	tctgtcctt	tttgtccttc	gacctgcata	tcccatctcc	92700
tctcccaccc	cgttcttatt	tctactgtag	ctagctgcgg	tttgtgatgt	gtgtaaccaa	92760
agacgcagaa	cagagaggaa	ggaaagggaag	cagtgataga	gttgggacaa	taagagaggg	92820
cggacccagg	agacctggag	aaatgggggc	actgtaccag	acttagtgca	atggcatcac	92880
agaagagggc	agaaccgagg	agtgggggga	agggaaaggca	acccatggga	ggcgggcttc	92940
aagggtggg	gaagtgatg	gatgcgaaat	agagaaaaga	gggacagaaa	agagacgaaa	93000
gccctggacc	ctccattaag	tgagaggggt	gggaagatgc	ctaaggccct	ttttctgtcc	93060

tgccttttcc	gattctgggt	ccctggggga	gctctggagg	tgaggggcca	ggaaaggcac	93120
aaggagaggc	ttgggtctgg	aggagagatg	ggtagccag	cagggtcac	cttccttcgc	93180
tgaaaggaac	tcctttgact	gtagctccct	gttttcatgc	tcagctgttt	ccttctcttc	93240
ctttgtgggtg	ccattcccag	ggcacaggct	atggaaacaa	aagccccacc	agcaaggcca	93300
aggactgtga	gccgaacctg	agactcagac	tggagggaat	agcatgggtga	atcccacatt	93360
ccaccgcact	ttggaatcac	cttttagcca	ctctgatgcc	caggttgcag	accagaccag	93420
ttaaatacaga	atgtctggag	gtgagagcca	ggcttccttt	tctaagatct	ctatgtgaat	93480
ctagtgtatc	taataagcag	caaagttag	gaagcatgaa	aagagttagg	caggcccagg	93540
ttcaaatccc	agctctgcct	cttcctagca	acagaaagat	ggctcagact	taaccttct	93600
gagctctcatt	ttttgcattt	agaaaatgga	gataaggata	tctcagagga	ttattgtgag	93660
gatgaaatca	gagagcacat	ggggtctgac	aattagtaag	tgagcagcaa	aggaatgccc	93720
ttcctctact	ccttgtggca	aatgactgca	aaaatgatca	catttcttca	cctcctctgt	93780
atttccccc	atttgaatga	gactgcagct	ctatttcccc	atgccctgaa	tctgggccag	93840
ccttgtgaac	tgcttcagcc	aaaagaatgc	agcagaagtgc	gctgtgccaa	ttccaagctt	93900
aaatctcaag	aacgcttgtg	catttctgca	ctctttcaga	accctgaaat	cacggtgtga	93960
atgagcccac	gtgggcttgc	tggaggatga	cagccacgtg	accagggcat	ccctgcact	94020
ccaaacctat	gtgagtggg	ccatcctagc	atagcaggcc	cccatgtaat	cctccaaatg	94080
atcagatgta	tgaatgagcc	ctgtcaaaat	catctacatc	tggccctgat	cagcggaact	94140
agccagctac	ccacagactt	gtgaaaaata	ataaatgctt	aacatttttag	gctgctgagt	94200
tttgagatag	tttgttatgc	agcaatagct	aacagatgca	ctgctccagt	cctcctcctc	94260
tcctgtgata	ggtttgcttt	accctgtcca	tcccacccta	gggccaatga	ggggctctgg	94320
cccacaatca	ccagatagtc	cttaccata	gctgtagtgt	gggggcagtgc	gggtatggat	94380
gtgcccccg	ggcatcagga	ggaaagtgcg	tgtctacatgc	caggtactgc	agacagagat	94440
gaagatgaag	gaggccctga	ggctgatgcc	tttttcataa	agaagctgca	gaaggagaag	94500
gaaaaagtca	gtgtcacacc	cacgttcata	gcagcactat	tcacaatagc	caaaggatgg	94560
aagcaaacta	agggtccatc	agcagatgaa	cagctaaaca	taatgtgatc	tatacacaca	94620
atggaatatt	attcagcctt	aaaaaaggaa	agaggcaacc	atgctggctc	acacctacaa	94680
tcccagcact	ttgggacgca	cgaggatcac	ctgagcccag	ttcaagacca	gccttgacaa	94740
catagtgaga	ccctcacccc	ttctctagaa	aatttttatt	taattagctg	gggtgtggtg	94800
catacacctg	tagtgcagc	tactcaggag	gctgagtggg	aggatttctt	gagcccagaa	94860
gtttgaggct	gcagtgaagc	atgactgggc	cactgcaccc	cagcctggac	aatgaaacat	94920
gaccttgcc	ccaaataaaa	aaaaaaaagg	aaaggaaaaga	aattctgaca	catgctgcaa	94980
catggatgaa	ctttaagagc	actatgcagg	gccaagctca	gtgggttcctg	cctgtaattc	95040
tagtgcttta	gaagaccaag	acaggaggat	tgcttgagtc	caggagcttg	agaccagcct	95100
gggaaacagc	aagacctcat	ctctactaaa	aataaataaa	taaatcagct	gggcgtgatg	95160
gtgcacgcct	gtaattccag	ctacttggga	ggctgaggtg	agaggatatg	attacatgat	95220
tacatgcctg	taatcccag	actttgggag	gctgaggcaa	gcagatcacc	tgaggccagg	95280
agttccagac	cagcctggcc	aacatgggtga	aaccccgctc	ctactgaaaa	tacaaaaatt	95340
agtggggcat	ggtggcacgc	acctgtaatc	ccagctactc	gggtgggttaa	ggcaggagaa	95400
tcgcttgaac	ccgggaggcg	gaggttgag	tgagccaaga	tcctgccacc	gcactccagc	95460
ctgggcaaca	gagcgagact	ctgtctcaaa	aaaaaaaaaa	ggttaagata	gtaaatttta	95520
tgttatgtat	attttattgc	atacaaaaac	atcagcagaa	gaggcagggg	ctggaaccct	95580
gttttctaa	gagtcctagt	acaagccatc	acctactatc	ctgtaagctg	attagggaca	95640
cctggtacac	acatgcccc	acccacccca	agacacaccc	ggcagttagag	gagtcctcat	95700
acgacccatc	cccacagccg	gtggagcctc	ctcgtgtggc	tccccagaga	tcttctagcc	95760
cagtgccttt	tttcccccaa	cgacagcaaa	ggccttttgt	tcaaagaaaa	ttttacacaa	95820
aaattcatct	tacaaaacac	accaatgggg	agcttgccag	tcctctccct	ctttattctc	95880
cttggtgact	ggtatgacat	caaagagaat	ccctaagttc	ctcaacagct	cagtttgaaa	95940
accaccgacc	tagcccaacc	tcctccatt	ttacagagag	tgacgttgag	gtccagagag	96000
gtgcagtga	ttgctcaata	aattgacaga	gtaagcagca	gcaaagttag	attaaactaa	96060
gaattcctgt	tcctgctccc	tttcccttcc	caactctaga	gagacaggag	agaggctggg	96120
catggtggct	catgcctgta	atttcagcac	tttgggaggt	caaggaaggc	ggattacttg	96180
aggtcaagag	ttcaagacca	gcctggccaa	catggcgaaa	ccccatctct	actaaaaata	96240
caaaaataag	ctgactgtgg	tggcacgcct	atagtcacag	ctactcagga	ggctgaggca	96300
ggagaattgc	ttaaaccac	taggcagaga	ttgcagttag	ccaagatccc	accaatgcac	96360
tccagctgg	gagacaaaagt	gagactccaa	ctcaataata	aaaaaaaaaa	aaaaagagag	96420
aggaaaagaa	gatgaggcag	ccatctgggt	tctccagggg	aaggaggagg	aaccagagaa	96480
gtgactctta	tgccaggagt	agaaaggctt	gagtgcctca	ggggctcagt	ctctgcataa	96540
ccctccaaac	ctccaaagct	tatgggacta	agctagactc	atgtctgggt	ggtgactgcc	96600
agagatcctc	ttctctgccc	ccataacctg	caggcagtgc	caactgcctg	tgacctaac	96660
ctaagcccag	agagaagtcc	caggttggat	ggcttgagat	ccacactctt	cccttccctt	96720
cactcagcca	ctgtggtgt	gctggcttta	gtcctccagc	ttgctgcctc	ataattgaag	96780
catggttgcc	acaactccag	ctatcacatc	ctcacaccac	aacattcaat	gaggaagact	96840
ttgtttttac	tctgctttca	ccttgcgtca	gggaagaaaa	gtcccttga	atcttccact	96900

atatacactc	ccttttatctc	attaaaaagg	actggatcat	atgctgacct	ccacctatca	96960
ctagcaacgg	gtaaatggat	tgccatgggt	ggctttaatc	aatcaggatt	catcccctgg	97020
gctaagcggg	tcactgcca	gataaaactg	ttcgcaatga	ataagacaga	atggttggtg	97080
attgacctct	aatagccttg	gcaacagttc	atccccgat	acccaacat	cagccactgg	97140
gacagctgga	caagcctctg	tgtctgcccc	tgtctgtacc	actagccact	tgccaccttc	97200
ttgtccaaac	tagaagctca	cagcagcaaa	cgccccactc	taaagggtccc	ccagcctcta	97260
cccaacactg	gcccgaagcac	attatgacca	ctgccacaaa	agcttgggca	agtctgaaga	97320
aggggcttag	cggttacaag	ctcaggctct	agaaccgaca	agcctggggt	caagttccag	97380
tatcatgggt	actagctgca	gaaccttcaa	caagcttttt	aacctcagag	actcaaatgc	97440
ttcatctgta	aaatgggggt	aacacagtac	ctacctcacc	gagttgatgg	agaaaataa	97500
tgcaggttca	caagacaagt	gtctggcata	tacaagtgcc	cagtgaatgt	aggctgttgc	97560
tatatattacc	ttaataataa	ggaagactgc	cgaggaagag	tcaaagtctc	cattgtacag	97620
agtgatgatg	gtcgaacggg	gttggccaaa	taggttccca	atctggggat	gataggacta	97680
gcctggatca	cttattttatt	catgaaacag	atacttctctg	agcaccagc	atgtggcaga	97740
ccctccttat	acccaaactc	accctccacc	gctagagctc	ccacctcagc	ttggggccaa	97800
cccctctgag	gcagccaatt	atagaaaagg	gtctctcctt	ccctccacct	tcggccacc	97860
ctgcccagtg	cctgggatta	gggaaggctc	ctacctcag	gttgggtgatg	agaaacagga	97920
ttcccccaat	ggtgagcatt	ggcatggcca	ggaagagcag	cacggctgag	cctggaccat	97980
caaagtcaga	ggtaggggtg	tgtcatagt	caacccaaac	atggagcccc	aaactctgct	98040
cccacctgct	ccaaattccc	aacaatcctg	gtatccaggc	cccaattcta	gccagcggtc	98100
cagcgtcctt	caagggtttt	taggataccg	gccaaggctt	cccagatat	ctctgtggaa	98160
gtcttctgag	ccaccttctt	cccaacccaa	gttggctcct	agtctgtggc	aggccaggaa	98220
gtctacagac	agaggcagag	ctctaagtga	agccacctct	ctcttccctc	agtaaaccac	98280
aagctgcctc	tccctttcat	ccttgacact	ctggaaaag	aagaccctgg	actcaggctc	98340
ctggctcaac	cctctagccc	attcccta	tcatgggtatt	ggccttgagc	ttcaatcatc	98400
tgtttaatgg	gaacaacagt	tctgtctctt	cctgtctcag	gtgctatgag	aactgagtga	98460
gaaaaggacc	atggtctttt	ctttgttcac	taaactctga	gcacttcttt	ggtgccaggc	98520
attgtgcttg	gcaactggaaa	tgcaagatga	atcagatagt	ccttgccctt	aaatagactg	98580
acatgcaaac	aaatggttat	aacaggtctg	gtaagtgtga	gaccacagca	aaaaagctca	98640
agagctgggc	taggggaacc	cttgacaaat	tcttctctcc	caaaccagac	ttctgcccac	98700
cattattctg	gccacaacct	atgcctgtcc	tattatttgc	taaaatgttt	taagttgact	98760
cactttttatc	caaaaagtat	ctatttttaa	aggacacttt	atatcactac	tgtagatgaa	98820
aacactggca	ttacttgtca	tgaatagaaa	gtaactgtca	aaataaatac	aatgaaagga	98880
aaacaatgtt	attcaattgt	agctggatgc	atgtgacctt	agaatgttca	aagcctaaga	98940
cctgctcttc	ccatcagtg	taaaatcaca	ctggccccac	atgaagacat	tctttcatga	99000
aatcagaagg	actgaaagag	aaataaaaaag	ggaatagctg	ttctaccagg	tgattttgatg	99060
tttgttagtg	tagttcacgt	agtatgcgtg	tgccccaaac	atcctcttaa	tcaccgtgct	99120
ataccttaag	aagcactgcc	aagagcta	tttagagtat	tcacacagtt	taccattcaa	99180
tttctgtctt	tataaaatgt	acatctctcc	tactactaaa	ggttggagac	tcctttcaca	99240
atagagtcct	tatgggctca	atgctttttt	caaaactgaa	aagccctata	ttatggaggga	99300
agaggaggat	tgttgctcag	acgatttgca	ggcacgagtc	aaacattacc	cagccaccac	99360
ctccacattc	agttgcttaa	aaatcattta	caggctttta	gagtagatga	tgctggtttg	99420
ataaggagag	tggtttgaaa	taattggttt	gagggtgctg	gccatctcat	gagatctgtg	99480
tgaacaaaga	cactcagcct	ctgtgtttgc	ccagctagag	tgacagacaat	ctcatgtgct	99540
tgtcagcttt	agcatagctt	acacacacaa	gagtaagtga	ctttctttcc	taaacccaaa	99600
attgagccac	gggtctaaca	ctaggaagga	atattgggag	gcattctcgtg	gccaccataa	99660
ccaaggcaat	gacagaaaga	agagtgaagg	atcaggaggc	ctgcacatca	ggccccacct	99720
ccacttgctt	tctctgtggc	catggacatg	tctttgcaag	gggtcctgct	gtggcttcag	99780
tttctcctct	gtgtaatggg	tggaagggtg	gtggaaaata	aaccagattg	gagttccaga	99840
cttaacagac	tggtgaaatt	ttaaaacaaa	gattttgagt	acaatagggt	tgtcaacttt	99900
taccctgcta	agtaaggata	tttgcaaaat	ggtcattcat	ataatcattt	cattaaaaag	99960
agaaagagaa	catttttaaca	cataggagaa	ggatgtaaag	gttttttgtt	gttgttttgt	100020
tttttagggg	tttttggttg	ttttttggta	gagtcctact	ttgtcaccac	ggctggagtg	100080
caatgggtgtg	atcttggtct	actgcaacct	ctgcctcctg	gggtcaggca	attctcctgc	100140
ctcagcctcc	tgagtggctg	ggattacagg	cgcgaccac	catccagcta	atthttgtat	100200
tttttttagt	agagaaggga	tttcaccatg	ttggccaggc	tggtcttgaa	ctcctgacct	100260
cagatgatcc	acccgcctca	gcctcccaaa	gtgtgggtg	tacaggagtg	agccactgca	100320
cctggccaga	tgtaaaagtt	tgaataaatt	ctactctctg	aagtaatccc	tctccatcat	100380
ccttgctttt	cacattttct	caataaactg	ttttcacaga	ccagcaatag	ctcaagatcc	100440
ttccaggatt	ctttcaagct	gcagatctct	gaataacat	gtgggtctgta	tatcttgctt	100500
atagccctct	gctcacacct	gccccagcca	ccagggtgct	ctgagcttgc	atccctccca	100560
cccacctgac	agcactcacc	tgcaagaggtg	aaggctatga	tgagtgtggc	ggtggtgtag	100620
aaaaactctga	aacacataac	aggaaagcca	gaattattgtc	aaggaggag	aaacctggga	100680
gaaaaaacat	gattctgctc	agccagccca	caagtgtagg	acttgaccgc	accctcagcc	100740

tgggatgcaa	cgggcactga	tgcctctgag	ccccaggctc	aaaaccaggc	gcaagaagcc	100800
gcgatgagat	tgagatgtgg	tcctgacctc	atggacagtg	catttttgctc	attctgaggc	100860
ccaaggctag	catggaaagt	cttggacaat	gagctcagct	gacgatgtga	ttggcttggg	100920
acttagccag	gacagaatgg	gcaaagcgaa	ggtcctccca	cctggaagcc	ccaacagccc	100980
aacccttgg	agaaaggggt	tagtgcctgg	tctgcaaate	aaggccttga	gttctaactc	101040
ctcctcactc	tgtgaccttg	ggcaaggcgc	tgtccttctc	tgggcctcag	gagccttttc	101100
tataaaaaa	aatgatcgga	ctgatctagc	tcagagtgtc	atgatttcag	gactacagtc	101160
ccaaggttat	caggctccct	tagcatttgg	gggtccttga	aggcatggag	taaaaaaaa	101220
aaagcaatat	cctaaggctg	gagaagaggg	aggggacaaa	ggaaggggag	gaaaggggag	101280
gtagcagggg	gccaaggacc	aagaaggact	gaggtacagt	cattctgcat	ccaaggctt	101340
aaattgttaag	ggactggctt	tactctggct	gtttccggaa	aggcaggccc	agccagccct	101400
cccgtctctc	tctctgacag	ccaatctcac	atgtgcctcc	ctgggagcac	ctgctctgag	101460
ctgtatcagc	ccccagcagg	ccgctgatta	ccactgagcc	tggccacaga	gcacgagatt	101520
aggatgcagc	aacacactgt	gtgtgagatc	acgtcccgaa	ccttctgact	catctgcaca	101580
gaaaaccccc	ccagtctccc	ctccagtcag	aagggacctg	aaattccacc	agtggcaata	101640
ccaagaatac	ttcctattag	ctaagccctt	agggagtgtg	tggctgttgg	ggcggtgagg	101700
gggggagggtg	aggaggtatga	ggatgaagcc	tgggcaacct	ggatgtgagg	ctgtgcaggg	101760
gatgaggaca	aggatccttg	gggtgaagga	agagaagagc	aatttttaggt	tttgctaatt	101820
ttgtaaccct	ggctccaagc	cagcccttac	aggaagtcac	cctggcctcc	ggctcaattc	101880
agcacgtgat	aggggaagcca	catttatgca	gagcagggaa	cgaggtaagg	aaatggaagt	101940
ggggctgttg	tgaagtgggc	aagtctagag	agagtccgc	tgcctggggc	tgttcctaac	102000
agctgctggg	agcgagctgc	aggtgtgggt	cctggcaggg	tggccgggct	gtctgactct	102060
ggatttctag	ccaagctagg	ctgctgcctg	aaggatttcc	cttaccacc	tttgctctgg	102120
ctggcctttg	ggactttacat	ggctatgagg	cgtgccacgg	tggctctgaa	ccggtcaaa	102180
atgtagccag	tggggaatgt	catgaagtgt	ttcatgaagg	acccaggggt	gaagatgagt	102240
gagaacctct	catcctgggc	tttgagctct	ggagtagaaa	aaaggtctcc	catgcatccc	102300
agccttctctg	ccaaatgagc	acacaggctg	ggctccccctc	cacctcagac	agcttgtcgg	102360
tcgcaaactt	gtcccttaag	ctgagttgaa	atgtggctgc	ccctaattac	ccctcaggag	102420
ctggtgcctc	cctcccaggc	acttcccaga	tcaagtgggg	tgagagctgc	tgacccttcc	102480
tctcatcata	gaaagagggg	tgggcagggg	cgagagtccct	tcctgctcct	tgccaccacg	102540
tgggagccag	acttaacttc	cttagaaaag	tcattccctgc	ccttaccagc	ctgccctgtg	102600
gcattgccaa	tcggcccagc	atctgggtcca	cacagatcct	taaagtaate	ttcattcttg	102660
aagacaaaca	ctagtgaagg	ccagccaaag	aggacgccag	caaagcccag	gcattccagc	102720
agcccagtc	gcagtgtggc	cacgtgcagg	ggcaggccct	ggcccgccat	gagcagaagt	102780
ggagtggatc	ttcaaateccc	actttgtcct	cctggacgga	tcacaggcgc	cgtaagcctg	102840
gcgttttgag	acttggaaaa	ttcctctggc	aagccaagcc	cttccttttc	cgtagctctc	102900
tggttgtttc	aggcctgggc	aaaaaccatc	agcgggtgat	tctctgggatc	ctgtagaata	102960
aagatagagg	ctgctggaag	aggaggcctg	cgggaaaggg	aaaggtagac	tagagttatt	103020
tgtgaggtgc	attaagaggc	aggatgatca	tggccgctgg	cagcaaatgt	ggggaataaa	103080
tactccaata	catcatctta	ggcactgcat	ttgagtaacc	acgtggcaag	tagagaggca	103140
ggtcttgatg	gccacctgga	gtcacagggtg	agataaacga	cttacccaat	agctccccgg	103200
gagcagggtg	agaagcggag	ctcctgcgct	cgaattctga	ataccgtccc	ctaaaataat	103260
gacagcaact	caaccagggtg	cagaggcagg	gagatttcat	acagaagaca	caaactcccg	103320
ctgccaagtt	ggtgttatct	tcagtttact	gacaatgaaa	caaaagctcc	catggatttc	103380
aggaaacttg	ccaaggtcac	agggctagtt	tagtcacgac	gcaggccatt	ctactgccag	103440
aaatacccc	aactcccatg	accctgcgct	aggactcgca	aacctgggtc	ccgcgcgect	103500
tcctcgcatc	aacttctacc	aggaaagcct	ccggggggccg	ctccccgcca	gcctcgcgac	103560
cccgtctccag	cctgcggcct	gcctcccccg	cagaggagcc	cgaggggcca	ggcgcgcgctc	103620
ggcgccccat	ggcgcccgaa	aggggacct	tcgcccatacc	cgctgctcc	cgccgggggc	103680
tctccgccc	ctttccgcac	gggcccagggt	cgcatctcg	cctctcgag	cccccccgag	103740
tcccctgctc	gcctccgccc	cctcctgccc	gcccgggaagg	ggttggggca	gacctccac	103800
tctccatcac	ttccttcttc	ttttcccttg	ctcacagcct	cccgcgcct	ttttacctct	103860
ccctcttgaa	acttctccct	ctagaacccc	ctagaacccc	agcgggtgtct	ttccctccct	103920
cctcgctgcc	tttcagcctc	ccagccccct	tgcctctgcc	tcccctaacc	aagttagttg	103980
aatgctgtta	ctcgctcagg	cccacctagg	gaaatgtca	caccagcac	ccagaggaca	104040
cacagacagg	acatgagggc	atagggacac	acacactcta	tttgtgcatt	ttgccttgac	104100
cgctggggtg	gcagggaaca	tatttttctc	atttgtctac	cagcttaacc	gtctctccca	104160
gtttcacact	cccagagctg	ccaaaaaat	cccaaccaca	gaatcaggaa	gccaagaacc	104220
aggactgagg	gcttttcaga	aaccatcccc	tggaggactg	ccccatattt	tcactcccaa	104280
aaacccctta	gatgactccc	tgcctcacc	ccgcccccca	ggttctgaaa	gagccttccc	104340
gccagactgc	attgattaac	cattcattgc	cccatttttt	attaatcaaa	gacatatata	104400
attgtcatc	ggagcttggtg	atcagcgtga	ggccttacta	agcagctgce	ttactatcct	104460
tccagcccag	agcacgtgag	ctgacgtctt	ctttggcctg	tgtggccggt	tccttgccaa	104520
aagctcagtt	tggggagagc	ttcttgcgta	ttagatgcag	tctgcagact	cccaacccca	104580

gctacctgga	tcccctgagg	gcccaggaa	tccagctatt	ccaagcccac	tcctcttttt	104640
ttaaagagga	agaaatagag	gttacgtag	gggacagcca	gaactgagga	ttttccagct	104700
caccacaaaa	gcacaaaaa	taaaagtctg	caaccaccct	agtgacttga	ctgaatggag	104760
gaagggtggc	tggggctcctg	tacccaagc	tactcactag	ttatacaacc	tgaggcaagc	104820
tcttttgctg	ccccacctgt	aagacgagga	caatagtacc	ttaattatag	gaattgtcat	104880
aaaagaagta	taagatgggt	gtatgagggt	cctgcatggc	gcagggtgcta	taggcagatt	104940
gtagggtagt	agattttctta	gtctgcagtt	atgtagacag	agccagagaa	gcagctctgg	105000
ggaggaatth	caaaggaact	tgcccacggt	cattctacaa	agctgcagta	ccttcccaac	105060
tctgaaacgt	atgctctcat	caccccgctt	taacaaacat	ttggacatta	gagaaaaaca	105120
gtcttttctt	aaaataacat	tatttatggg	agaaaatcca	caaaaatata	gcatcccagg	105180
acaaacaggg	cttaagatgc	aagattttct	atthttactgc	aagacacaaa	gactctgaaa	105240
ttaatgcatg	ccctatcttc	tgctctggca	tacatttttag	tctcctgggg	ggatcagtaa	105300
gtgtggaagt	agcaaggag	aaacagaaaa	aagtcaaaagt	aaagagacag	attttagaat	105360
gttaatctgc	aggagcctgc	cagaaagatc	tagctcatgg	gctatctgta	catccaggac	105420
tgaagcacgg	gacacggggc	aggctcgtcca	gggttctgtc	caccttatct	tggtacctct	105480
cttgactctt	agagcctcca	ctccacatct	ccatcgaatg	tctgcagaag	acgtggcctc	105540
actaacaaca	agtcttactg	aactgatggg	acaggaatth	agaatatcct	ctgaaccatt	105600
cccatgttct	ttgggttcgaa	ttccagcagc	tagaaaaggc	agatgctatt	ctgatcactc	105660
tcctgcgtgg	ctccaatgag	gattaatgag	taacatcaga	gagagaagtg	attataataa	105720
ggctctgacg	tgcacccgat	gtcttcatcc	ttttctcttc	gcctccttcc	tcatcatctc	105780
acaccttttt	ttttttaatt	gactgattgg	ttcaacaaat	acatgtggta	cctcaggctc	105840
tgtgccaaagt	gccgggattc	gtagagaaga	gattcagtgc	ctgctctcaa	ggggctcatt	105900
ctcttgtggg	agagacagac	aaagaaaccc	aagattttctg	gagtgtggga	atggctcttc	105960
aggcagatgc	tagcacagca	cattgaaagg	cacggaacct	caacaaaaca	ataacattta	106020
ggaaccagct	agagcacagg	gtggtgaaga	aagtggaaag	atttgaggcc	agcgtcgcca	106080
tctaagttag	ggcattaaga	attcagccca	catcaatcaa	tcatgtccta	ttgatttcac	106140
cccttaatat	ctctcctatc	tatccgtggc	cactgctcta	tgacagacact	catcatctct	106200
cacagaggca	tcactctgctt	ccaagccatc	gccattctct	tgcaagagtt	tatttccatg	106260
gttcccactg	gatggcttca	cttaactgct	caaaacccct	ctgagggtcca	gtcaactggc	106320
tggttaaggac	cagctccagg	tctggggatg	ccagccatga	gacatttgctt	tgaggggaag	106380
aggagacata	gaactggatc	tcctgcatcc	tactgcccac	gtaccaatgc	tggagggtgg	106440
tttcttctcc	atcatcagca	agtctggata	tccaggatcc	accctatgga	tggttttatg	106500
gacagagtgg	gaagatggat	atgttttaggt	tagggaaaga	gggtttgcca	aagaggggcag	106560
tataagttag	ctgcactcca	tcattccctt	ggcacaacaa	atggctagta	tcctctagtc	106620
ctcaagagca	ccaccttcca	atgcagctcc	tgctgttcca	cagacctctc	tcctcaaact	106680
tcctctgaac	aacctcagcg	agggcaattg	ccactctctg	ggcagagtcc	agatattctc	106740
tcctaccctc	tgacatcact	ttctaatttt	gtatatgtag	ataaactctg	agccattcac	106800
ataaagggtct	ttgatttctg	atacgccaaa	cacataaaca	aacaaacata	agctttctct	106860
tcacaatggg	ctcatgtaaa	ttaaaatgtt	tggttttcca	cctacgggtct	tgaaaggggt	106920
ttctacagcc	tggttttgaa	gtcagaaagc	aaaaggtaaa	tgcaaacatc	atttcacctg	106980
cagagaaaaat	tctaatectc	ttgaggcagt	gccaaaaata	atacaagcac	actgctatcg	107040
agccaattac	tggtatctct	gagcttccgt	ctcctcatct	ataaaattgg	aatcgagctg	107100
tatggatttaa	agataatgta	tgaaaactgc	ccaatttagta	cactattaat	aaatagcagc	107160
tactgttggt	aacaaatatt	attgacttac	tggaaaaaca	agaggaaata	aagtcacatt	107220
tagggagaat	ttcaaagtgt	tcctaacctc	aaaaagaaat	aaattagggg	gaaaaccact	107280
aagtaatggg	tgagctcagt	ttaccttgct	taagaagtcc	caccctagag	aactgatctc	107340
tagatgacac	ccaaatgcac	tcagtacaac	cccccaagac	tgtctgggct	taaggcaggg	107400
gcttggtattg	tcctgtaagc	tgtgggaagt	ctgttcatga	gccacagtag	acaggaaggg	107460
gatggagtct	tagagctggc	tcttcagggt	atctcctagt	gtgttcaaag	cagttctcag	107520
gagggtgggg	aactactaca	tagccaagta	aatatgaggc	ctccttgctc	tggggagacc	107580
tttctcttta	acagagggtga	atctgaaagg	atacccaaag	aggcactgga	gggtgggggc	107640
cactctggcc	cctcagagca	gccagctcag	cttcagtggg	tgctagaggc	agcagaggat	107700
cagcctggat	cagcctccct	ttcacatgc	agaaaacaga	gctccccac	cagaccagat	107760
actggaagca	ctggggccagg	cctaagagaa	agcagagccc	caagccccac	cacaccaggg	107820
cttatgaggc	tactgctacc	caccttccca	agccccagct	ccacttctat	gttcatcaag	107880
caactgttta	ctggttaactg	cacttcccg	tagactttgc	tagaaaggaa	tgctctcag	107940
actgacaat	atctaaacct	gcaactctaa	ggactaggct	ggggaacact	gtgtcaacat	108000
ggaggcacgt	cctacccctg	agaagaaaaa	taaggaatat	caataatacc	tgctgggcac	108060
tgagcactga	ctatgtatct	gattcgaagc	gctttttgct	taatctgtca	ctgaatctca	108120
cgataggtgt	tggtattagc	atctattatc	tgggccaact	gaggcctaga	gggacgaagc	108180
aactccccca	acatcaccag	gtagcagtg	caggactggg	atagaaacct	gatgctctga	108240
ctgaaaactaa	tgcttttttt	tttttttttt	tttttttttt	tgagacagca	ctcactctg	108300
tcaccaaggc	tggtgtgaaa	tggtgtgata	tcagttcact	gcagcctcca	cttcccaggt	108360
tcaagtgatt	ctcctgcctc	agcctcccga	gtagctggga	ttacaggcgt	gcaccaccat	108420

gcccagctaa	tttttttgca	ttttttgtag	agatgggggt	tgcctatgtt	ggccagactg	108480
gtctcaaaact	cctggggtca	agtgttctgc	ctgccttggc	cccacaaagt	gctaggatta	108540
caggcgtgag	ccaccatgcc	cagccagctg	atgctcttaa	tctgtgccct	accagcctt	108600
cctgggaggg	ttcccaagag	ctacacagag	catgagttct	ggaatcgggt	tgatgggggt	108660
accagttatt	actaatagga	atgaagatgg	gtaattcttt	cagacagcac	ccttgattaa	108720
aacaagagag	tagtgctgcc	tctctgtgat	tctgtgtctc	cctgccctgc	tcacacagac	108780
accacaccca	cccacacgca	tgatcatgaa	aagaggaaat	ggatccagga	gaaggagacg	108840
actcctgagt	gaaaacaacg	gggtttttca	cattgagagc	tttgcccaac	accccaaaga	108900
tgaaaagagc	aggaaactgc	tggggccgat	tgaacactgg	acttttgttg	tggaaaaagg	108960
caaagggaag	ccggaagaga	ctggaacagt	ttccatggtg	ctggaggatg	gggaagtggg	109020
tagggattag	ctggagggag	aaggagaagc	tggggtggga	ggggaacctc	cacttgccag	109080
gagagcacat	gtaggatggg	aacccagat	gatactcaag	gcatggcatt	agaccagaag	109140
caagtctgtg	gtgaaattag	ggaaggctcc	actgcggact	gtagacagag	cactggacaa	109200
ggaagtggga	gacccagggg	ccagtcctgg	ctctggagcc	ccctgggtgg	gctgccccgg	109260
gcacctttct	ctcttggggc	ctccattcct	acctctgtga	agcgagtgtc	gaacctctct	109320
tagccctgac	ttgctgaaat	gctgggacac	tgtacactgg	ctgacattaa	caggggattc	109380
gtcgtggggg	gctgcaatgt	tcctccagat	gctgcacggg	agagggcaga	aaaggcctat	109440
atggtgagtc	cgccctggga	gcctctgctt	ggaagctgaa	gtggcctgag	agtgactcag	109500
aaaccacgga	agttcccggg	gctgatgggt	tcttatagat	tgtacatgca	gctctcctcg	109560
tgggctgcaa	aaccgcaaga	tgggctgtga	ccactctcaa	ggaaagagcc	ctatctgcaa	109620
aaagcattct	gccctccagg	tcttaaagca	aacacagact	caatccttat	tccttttaag	109680
acaaaattgc	ctcaggggca	tcagggaggc	atcaggcctc	aaatgtgtgc	ctttctagaa	109740
ttctcaatga	aagcacctct	ttgggtatta	ataatgacaa	cagtaatgac	agtcatttac	109800
tgagtgtctg	ttttgggaca	ggcattgggc	taagagctat	atgtaatata	tattattatt	109860
tgatgcccac	agccacccca	taaggaggcg	gaggtactat	cattatgcca	actttaaaga	109920
tgaagaaact	gaggcctcaa	gagatgaagt	aacttgccca	agtcactcag	ccagtaaagt	109980
ggaagagata	gacttcccag	tatccagagc	ccatgttttc	accattatgc	tgaagtacct	110040
cttttcctgt	gccaatgtga	tctgcctcca	ggaatcctgt	cttgatgttc	ccttccccat	110100
acagaagtcc	tctctgtgtc	ctcttcagcc	tgatagtata	tcttttcata	ccattctttg	110160
gacatctctg	ttatactact	ccaatgggtg	tcctctccct	acccctccct	ggagagctag	110220
ttgttgtgat	taagtatagg	ggaaatgacc	cacactaaac	aaactcataa	gagactgatt	110280
gataaacctg	aatgcaatt	tatttaattaa	cactgagaaa	tgaaccacc	cagcagatgg	110340
gaatcctaag	gctgactggg	cagcaccaatc	tctttcagga	aggacaggct	tttgggaaag	110400
gaaatcaata	ccagaagggt	ctttgttgag	tacaaagtca	gagggaaggg	agttgatgga	110460
ttgacacata	ggtgaagctt	gacatacctc	tataaagcct	ccatcctgcc	aaggatcaga	110520
atatccaagg	caggagacca	tctgggtgtc	ctctcctttg	gacagtgtcg	gatttttctg	110580
gatcctatga	agatcttacc	ttcttggtcg	catttatcat	gattgtggaa	ggctttttgt	110640
ttccttgttt	gcttagatta	atttctgcgt	atttaataga	actgaaaggc	aatttcccat	110700
tgagacccac	tgaagaggaa	taatcaatac	atactagtgt	tggtgccctt	tgcagagaat	110760
tcacttctgt	gttgctactg	tatcctcatg	cttccttata	atggaggggac	agagatggta	110820
aaaacatgga	cttggaagcc	agaccgtctg	ggtttgaatc	ctggctctgt	tacttataag	110880
ctctgcaacc	tcgggcagat	tacctaagtc	agttttccct	tctctgaatt	ggggatataa	110940
tagcaccac	ctcaacatct	gtcaagagga	ttcaatgagg	gaatacacat	aaagtatctc	111000
gaacagtgtc	tgccatctgg	taagcagtcn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111120
nnnnnnnnnc	taatttttgt	atttttggtg	gagacagggt	tttgccatgt	tggccaagct	111180
ggtctcaaac	tcccgcacct	aggtgatcca	cccgcctcgg	cctcccaaag	tgctgggatt	111240
acaggcatga	gccaccacgc	caggccccac	acacctttta	aacaaccaga	tttcattcat	111300
cgggagtgct	ctgtggggct	ggtgtggaca	tgtgggtgaa	ggtggcactg	ggagaagtta	111360
ggattctcca	tgacctctgt	tactcatatt	cccacactcc	tcaaatttagc	ctgagtctcg	111420
aggacagtct	gatggctggg	caaaccctgc	ggcaaaccat	tccccagccc	tgccctctca	111480
accagagtcc	ttccgataca	tgattctggg	cagctgttgt	taccctgtgc	ctccatgttc	111540
ttccagagat	atccatgcat	gcatcggcat	atgtgtataa	ttatttatatc	tattatttcat	111600
cccacaagct	tttgacatca	atagtagcat	attattaaat	tgttctgtac	attattttat	111660
taactttgta	tctctggtac	tgctcaatat	aagaacatat	agacctggcc	aggcacagtg	111720
gctcacatct	gtaatcccag	cattttggga	ggctgagatg	ggtggatcac	ttgaggtcag	111780
gagttcgaaa	ccagcctggc	catcatgggtg	aaaccccat	ttctactaaa	aatacaaaaa	111840
ttagccaggt	gtgggtggcag	gcgcctgtaa	tcccagctac	ttgggaggct	gaggcaggag	111900
aattgcttga	acctgggagg	cagatgtttgc	agtaagccga	gatcacgcca	ctgcaactca	111960
gggtaggcaa	caaagcgaga	ctctgtctca	ataaaaaaaa	aaaagggaatg	tatagacctt	112020
ctttattctt	tttgatggct	gtaggtggat	gtttctaaaat	ttgtgtaacc	aatctcctat	112080
tgataatatt	taagttatgt	cttcagcatc	atatgaaact	tacaaacaag	gttgcatgga	112140
ctatccatct	gtaaatgtct	tttgaacat	attgcaggat	aaactcctaa	taactcctaa	112200
aatgagaatt	tctgggtcaa	agaggatatg	cattttacat	ttaatagata	tttgtcaaat	112260

tgtcttccaa	agtggtcgta	ccaattaaca	ccccgacctg	taatgaatga	gagtgccttt	112320
tttccccaca	ccctggagag	atgaaaaatt	tatgggccc	ctttggagtg	catgggtggag	112380
gaagctgttg	gccgttatat	aaccctcgtc	attaataagc	ctgggggtgg	ggggggagaa	112440
agagaggtta	gttagtggtg	gcaaacatag	aattagatag	aagtaataag	ttctaattgtt	112500
cgatagcaga	ggagggtgac	tatagttaac	aacaatgtat	tgtatatattc	aaaatagcta	112560
gaatggagga	cttaaaatat	tccaacacat	agaaataata	aatgcttgct	tgcggccatg	112620
ccaccctgaa	tgtgccagat	cttgtttgtt	cttggaagct	aagcagggtt	gaacctgggt	112680
agtattttgga	tgggagaaat	gataaatgct	tgagggtgata	gatataccta	ataccctgtc	112740
gaacattata	cattctatat	atgtaacaaa	atatcacacg	tatcccataa	atatgtacaa	112800
atataatgta	tcagtaaaga	gagggctggg	cacgggtggg	cacatctgta	atcccagcaa	112860
tttggcaggc	cagggtggga	ggatagcttg	agggcaggag	ttcaagatca	gcctgggcaa	112920
catagcaga	ctctgtctcc	acaaaaataa	aaaataaaaa	cgaattagcc	aggcgtgggtg	112980
atgcatgctt	atagtcccag	ctacttggga	ggctgatgca	ggaggattgc	ttgagcccag	113040
gagtttgagg	ctgcagttag	cctacgactg	caccactgca	ctctccagcc	taggcaacag	113100
aggaagacca	tgttttctaaa	agaaataaat	taaataaaaat	aaataaaaaat	aaaaagactg	113160
aaaagcagag	tggtaagaga	aaggactttg	gggctcaaca	gtactagcct	tgaacccttg	113220
ctgttactta	cccatcggtg	gataagcaaa	tgctttaacc	cctgtgtgcc	tcactttctt	113280
aacataataa	atagaagtaa	aaatcatacc	cacttcaagg	gtcattataa	aaagccaata	113340
gagataatgt	atataaagct	tctggaataa	tgcttggcac	acagtaggag	tttaataact	113400
ggaaattcat	tgttgtagtg	ggcagccttc	tgaatctgtg	tcctctttgt	ccactaatgg	113460
ctttgatctg	gatttggctc	aggcaagacc	tggggaagg	cagagactga	gggcaactgg	113520
aggtataggg	tggctctgagc	ttccccagca	gagtgaggct	gggaaaggct	tgggagacag	113580
accaggcagg	tgctgataag	accggaatgg	gaggctggag	cataaggcag	ttcagttttt	113640
cccaaagggg	ggtgtacaaa	acgatctcgt	atgactcctt	tatactgtta	atgtttttct	113700
tttatcgcg	actgaaaaac	aaaaccaaca	tatttaatga	atgattccaa	ggggattcct	113760
gcttttacaa	aaaatgctaa	agtaggcatt	cacatgttta	aaaattgagt	tgattttaaat	113820
tttaaaatta	ctaagtcata	gtacataatg	tgtgagccac	agctatcccc	aaaatcatga	113880
tagcgataca	ttaatgactg	aagttcttta	aacatcaaca	tacaatgcca	attccagaat	113940
tcagctcaaa	ttctgcaatt	acacaggctg	gggttgaaac	ccagcttttt	tgctaaactgt	114000
gtaaaattag	gcaggagggc	taacctcgct	gaatctcaga	tgtctagtct	gtaaaattgaa	114060
gataatgttt	gtttttatct	cacagagttg	ttgtgaagat	tcaataaaaat	cacaactagt	114120
gaggatgatc	tggctgtgac	acctgtcacc	ccactgatct	ccagagttga	ttcggtgat	114180
caggctggct	gggcagggtg	cccccttctc	cctcaccact	ccgcatgcat	tcctcccgaa	114240
actgcacact	tgggtcaaaga	ggaagacctt	tcctgataga	ggaggaccat	tcctcagtca	114300
aggttatatg	agcacctgtt	ctgtcctgcc	agaatctccg	aaggagctct	cagtaaaatc	114360
acaagatttt	attgtgcatg	gtagcatgag	cctgtaatcc	cagctactca	ggaggccgag	114420
ctaggaggac	tgcttgagcc	caggagtttg	agaccagcct	gcgcaacata	gtgagaccct	114480
gtctcaaaaa	aaagaaaagaa	agaaagaaaa	gaataataat	agtaataaat	cacctgtgca	114540
acgtgctcac	ttctctcttt	ggaatgtagt	aagtgtacct	aataaatgtg	atcattgtaa	114600
tcatacacagt	gagcacaggc	taaagcatct	tgactttatt	ctataagcaa	taaaagagga	114660
tttgttttta	cagaactcat	tatgttgtga	aaataatttt	ccaacattaa	caaagaacat	114720
tcttcaagta	aaaggaaaac	cacccatcat	tctcccaacc	ttcaataaatt	ttcaattttg	114780
catattctcc	agactttgtc	aacatgaata	cttactttac	atggctcgaa	tcagtggtca	114840
tgcaaattct	tttatcctga	cttttataaa	caaataatgat	gttataaacc	ggtctccatg	114900
ttcttgcata	ttctttataa	ttatcatttt	gtggctgcat	aatattgcat	tgactatgtt	114960
aaactgcagt	ttcttaacca	tttactgttc	tggggaaatg	gaggataatg	ccagggtcat	115020
gcctggagct	tttttttgct	tattgcatga	tattcttaag	atcaaattcc	agcagtgaga	115080
ttagtcagtc	aaaaagtaat	aatattttca	aggctcttgt	tatatatttac	tagattgttt	115140
tccagagttt	tgcacactgc	tcccagagat	gtaggaacac	agacgtcatc	caaccttgcc	115200
agtgtggtg	gatggtgttt	ataaacttct	gctaatttaa	taagtatgaa	atgctatcct	115260
cacacggctt	tcattttctat	ctctttgatc	attaacagg	tgaactattt	tccaagtatt	115320
tgtttactct	ctgcataccc	tcttgggtga	agtagtcac	cacttccttc	acctgtttat	115380
ctgttggaagt	cttgaggctt	gttttataaa	tgtgagcgag	cacttcagag	tcaatagaca	115440
ttaattgctt	ccagccagat	ttggccactg	aggctcctga	gcaggggaa	gcatgatcaa	115500
aactacaccc	tggacagatt	aaatttaattg	gagaaaatgg	gctgagaggc	agagatatgt	115560
gtcactggcc	tactgtgttt	gatcctatag	tggggggcctg	aactggggca	acggcctgag	115620
tccccacta	ccagtagcag	gaggctccat	gtgtccccc	tattagagct	tgcggcactt	115680
ccatttgccc	cacctctac	aataccccac	atacatgtac	tcactctccc	ttgcaaatct	115740
agtggcttca	ccccacagaa	tttaagggga	aaggaattgt	tctgtcctgt	tcacttactg	115800
cagaaatgag	aaaagcggtg	ttcacatggg	atcaccta	gaagggatgc	catccccaac	115860
ggtgcctata	aggaaatggg	ggagggttgg	agagttgtgc	aaaatgcaac	agggaatcat	115920
cagagtctct	tgccccatga	tagagggttc	tcaaatataag	agagtctaca	gcaactaatc	115980
tcacggccac	tctaggcagg	gcttcccaat	gcttccccaa	ccccacctcc	atcctagact	116040
ttaccacctc	tgctgaacac	agatgtttac	catagcacct	tgcaccatga	ttgttttgatt	116100

agcacctccc	acagtagact	gtgttttctga	taggtcagca	acatttgctg	agcacctact	116160
ctgcagggt	gtgccagggtg	cacaaaaataa	acaaagccaa	agacaacatg	gaccctgaac	116220
tcagcaagtt	cagagtcaag	tgggataggg	aggctctctt	cactggaagg	taactccaag	116280
aaacaattggg	actcaacttt	ctaaccaaga	gaactccagg	gagctaaaat	tctgacttct	116340
ggttaagact	ggtgtggagc	ttcattaaag	aagaaaagat	tcaccagagc	ttgagttcat	116400
agcctggctt	tgcagctttt	aagtcatgta	acctttgatg	aagttatgtg	acctctccac	116460
ccagctgccc	ctaacacctt	gcaggggagc	ggctggagtg	caaagggagg	cactgggtacc	116520
acagcctggg	aggcaccacc	ccactagtgc	aagccgggca	acctctgccc	ccaaggcatc	116580
cctagcctcc	caactgcaag	catcaatctt	gcacttggaa	aggaacctca	cctttgaaat	116640
ctaggttcaa	atttagaatg	atccagctcc	ttgaagtctt	atacagaaat	acagccagca	116700
gccaggcccg	gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gtgggtggat	116760
cacttgagag	cagaagttcg	agaccagcct	gaccaacatg	gtgaaacccc	gtttctacta	116820
aaaatacaaa	attagccagg	catggtggta	cacgcctgta	atcccagcta	cttgggaggc	116880
tgaggcagga	gaattgcttg	aaccaggag	gcagaggttg	aagttagcca	agatcgtgcc	116940
attgcactcc	agcctgggca	acaagagcga	aactccatct	caaaaaaaaa	caaaacacac	117000
gcagcttccc	tccacttccc	aaccacagct	ccatctcaga	caacaagggg	cctcatgtcc	117060
atgcacatgc	atatccaacc	aacatgtcta	aggcccaacc	acacctctc	caaacctctg	117120
ccccttggcc	accctttggc	catgggttca	tgcaactggc	gaaaggtagt	tcagagaaga	117180
agcccaaaaa	gggcccggaa	gtccacttgg	gctttttgag	attccagggt	ccaggataac	117240
ctaagtgtgg	tctagaagag	agatgcagct	tctgggaggc	acattccttg	gtcttaggga	117300
cttcttggcc	ccatggaggg	aaactggcta	gatgagggcc	aaagcagagc	cctctaaagc	117360
acagggtcca	gggaaggact	ctttttgacc	agatctaaga	gcagcactac	ctctctgagc	117420
ctgtttctcc	atctgtaaga	aggggacatt	aatagactct	ccccgctaga	gttactctac	117480
atacggcagc	acacgtaagt	tcatgacatg	aagcaagggc	ttaatatata	cccgttgtac	117540
tataaataat	aggccaggcg	tgttggtcca	cacttgtaac	cccagcactt	tgggaggccg	117600
aggaggatgg	atcacaaagt	caggagtttg	agatcagcct	ggtgaaacct	catctctacc	117660
aaaaatacaa	aaattggccg	ggtgtggtgg	cgtgcacctg	tagtcccagc	tacttgggag	117720
gctgaggcag	gagaattgct	tgaacccggg	aggcagagtt	tgcaagtgagc	caagatctca	117780
ccattgcatt	ccagcctggg	tgacagagca	agactgcata	tcaaaataaa	taaatacaca	117840
catacataca	tacatacata	catacataca	tacatacata	catacaatac	atggacaggg	117900
accctaaaaa	tgagacaggg	aaagagaaaa	acatgttctg	acaaccttgc	cctttatact	117960
aatttagggt	ttcttgccctg	ttttagaaaag	ggccttgaca	ggagccctgt	tccctcagg	118020
ccaggcagaa	caagggtgtg	aactcactgt	ggaagggttc	tgggtgacaa	gtgcagcccc	118080
gtccctccac	ctcccagcac	agtaggcagc	acgtgtctcc	attgactggc	tcaggagcag	118140
gcctggtgac	cagtgggaga	gctgaggagc	ccagggtggg	gtctgaagga	atccctagaa	118200
aatctgattt	tccccagggg	cccacatcac	gtgccagag	ctgggaaagt	ggaggcagca	118260
tgggatctag	ctgagaggct	ccatttttgg	tagcttctag	tttgggagtc	acagagacac	118320
cttgatgata	cgaagatgta	gctttgcagg	actctctaga	acatggagtc	caagatattc	118380
ccttcaatga	tgggacactg	aagccacag	aggagaggtc	tgtcccagtt	actcagccat	118440
tcggaggcag	agaccaggct	agaactcagg	acttttaatt	tggaccagga	ttccttttac	118500
cacagtgggc	agccctagca	agtgccagg	agggtggaac	tgtgaaggtc	atccgagggg	118560
tagtacacgt	gggtaggaag	tcatatctaa	gaactgaccc	ccagacctgg	ctctgccact	118620
cactccttat	gagaccacag	gtgctgggtg	cagtgggtca	cacctgcaat	cccagcactt	118680
tgggaggcca	agcgaggcag	attgcttgat	tcaggaggtt	cgagaccagc	ctgggaacaa	118740
tagtgagacc	cccacctcta	ccaaaattag	ccaggcggtg	tgggtgtctgc	ctgtagtccc	118800
agctacttgg	gaggctgagg	tgggaggact	gcttgagcct	gggaggcgga	ggttgcggtg	118860
agccaggatc	atgccactgc	acaccagcct	ggatgacaga	gtgagacaga	atgacacact	118920
gtctcaaaat	aaataaataa	atgacagcag	atcatcattt	ttctttctgc	ctctagactg	118980
caatgcctat	ttctccagg	agtcactagg	ataaaagtaa	aaataaatatt	atcagcattt	119040
accaaataca	gggtcagcta	ctctgttatg	ttctttcatg	ctttgtttct	tttaagcctc	119100
aaacaactct	atgagctggg	aacaagtatc	gtccttcttc	ctcccatctt	atttatttat	119160
ttatctatgt	atztatctat	ctattcattt	atttatttat	tttgagacaa	ggtccttcta	119220
agtcaccagg	gatggcctca	aacttgagct	aggaactagt	gtcaccaccc	cccaatttct	119280
tttattgatt	gattgattga	ttgattgact	ggttaatttt	gaggcagggg	tctcgctaag	119340
tcgacagggc	tgggtcttgaa	ctcctagtct	taagcaatcc	gcccgcctca	gcctcccaaa	119400
ttgctgggat	taccaacacg	agccaccatg	cccagcccc	cccatcttct	gaataggaaa	119460
actgggggtt	gaaaaggtaa	gcgacttgc	caaggtcccc	tccctagcta	gagagcttca	119520
gagccagggc	caaaacccat	caaagcctgt	gctctcgccc	attgagccac	cggaacctcg	119580
acactaaccg	ccaagtgttc	tacacagtga	aggtgacaaa	gaggtgaagg	gaagagccag	119640
ggaggttctg	tggactcact	cggtgggtat	gcccagaggg	aagggggatc	ttgggtggca	119700
cattgagagt	agctgcgctg	ttagtaagtg	agaactcgga	agtccagact	catccagtct	119760
gtgccaataa	acccctcctt	ctacctggtc	tcctttccaa	agccagctgt	tctccagaca	119820
atgggggtgg	cggggggcgg	tgtcctcctc	cttctcaggg	aaaatccgac	gctgagccca	119880
tctccagaga	tcttggcttc	ccgtggggct	gcagatccac	ctagagccac	cagagggcgg	119940

gccagcactg	cggccaaaggc	ttgaagaccc	agcacacca	agccccggcca	agcctccagc	120000
ccagtgtcca	agagtccagc	cagaggccga	gtcctcgatc	tcaaaatgtc	taactgcaga	120060
agcccaactc	atgttcaggc	atgatgtgtc	tgattctact	gggacaatca	ttgccacca	120120
agaattactg	ccaaaatagt	aacgacatta	gctacctacc	acccctccac	acaccaacac	120180
acctcatttt	accaagcact	ttctcatgcc	tgggggtgct	ggagacttaa	tgcagcctcg	120240
cgatgaccgg	gtagcctcac	cgtacagatg	aggaaactga	ggcacaagga	aggggagtag	120300
attgtctagg	gtcacctgga	gaactctgat	ctccagactc	aaatttccaa	ttcgtccccc	120360
ctccccccaa	ccctaaccgg	agctaggtgg	ggtggggaca	gcaaatgtgg	atggggggag	120420
gtaagagggg	tcagagtgtc	ctacagagaa	gaccaaatgc	attgtggcac	ctactgtaaa	120480
atgagaccay	ccagccaccc	acccaccagc	cagccaccta	aaagtcttca	gtgggcaoct	120540
gctggaaaca	cgacaatgga	tgacacgagt	tccttgcctt	caaaaagctg	gtagtctagt	120600
tgggggtgga	gggggtgagt	cagcagataa	ttatgggaaa	ccgtgacacc	tgtataaggg	120660
gcggggatga	gcagaggggc	tgcgactgcc	tggagccagg	gattcccggg	cggggcttcc	120720
ctttcctcgc	agctcgtccc	aggaggagga	gctccccccc	agcttcgggg	ttccgcctgc	120780
cttggggggc	cggggtcccc	tcccacccct	ccccgaagag	cgcgggcccc	gggaaccgat	120840
gacagcacac	ctgagtcagc	ccgcgcgcca	ccgcgcctcc	agcgtctgtc	tccgcatctt	120900
gtgatatttc	gctccccggg	agccagcccc	actgcgtccc	ggaggcagct	cggcaacaaa	120960
accagcgac	agattgtgcc	gcggctcatt	ccggggaagg	acgccaaacc	ccaccctgct	121020
acccccaa	ctccctcccc	gccgcgcctt	ccaggccctc	cccccaggcg	caggccctag	121080
tcgggggtgg	tccgtgggaa	acgcagggtc	ctgtcctgcc	tcctggaaat	agggggagcc	121140
ctgggtgagg	gaagacggga	gccccagaga	cttttctttc	tgtttctacc	tgatccgaaa	121200
acgagagggg	cgggaaagga	aatttagggg	cacagagagg	agctgggggg	ccgagaaggt	121260
ccgaaaaatg	aaccagcagg	gggcacccga	aggccgaggt	gcccacgggc	cgggagcctg	121320
ggaatgaac	tggggaagag	ggggagagaa	agggagcgag	agacaccgag	acacacagag	121380
acgagagaca	gagacgcagg	gagccccgcg	gggaggagga	gagagacgaa	gacacagaga	121440
gacagtgaga	aagacagaag	accgggcagg	gaaacagacg	agtagagaca	gaaaagggtc	121500
gagagagagt	gagggaggga	gggaacagag	agacagagac	cacgaaatat	gagtaagagt	121560
cgggggagaa	aaccagagaa	atcgaatgag	aacgcgagaa	gaacgagaga	ccgtggaggg	121620
agcagagaat	gaatgggaag	aataagacca	acatttatca	agagccgact	gtatgccagg	121680
cactgcattg	gaccctggca	cggataggaa	aggaggagcc	gcggcgcgcg	cagcggggcg	121740
aggggcttct	gtgctcgcg	gagcggcagc	ccagggggct	cagcagcccc	ggcaccgccc	121800
caactgcggc	tcacgcagcc	ccaaccccg	cagcgtgcc	tggccaccgt	acccgaagcg	121860
gctcccccca	gggccccgag	cctatcctac	gccggggcg	ctccgcggac	gcgcggggcc	121920
gagtcgaagc	ctgggtcaac	cgccctgcag	cctttgtagg	gaagtgccta	ggtgatgggt	121980
ctgctgatac	cgcctgtgac	caggccatga	agggccagag	gggctccagt	gagaccataa	122040
tccgccccct	tttaaaagg	ggtagaggaa	gttcacgcga	agccaacagt	cttctcccca	122100
gctttgggtc	ctctcctgca	cccccgcg	agataaagtc	tccccctccg	gacacatcat	122160
acatacacaa	aaaaacgcac	acactcgcac	gcgcgcccac	ctcgcacccg	cttgtaaagt	122220
cactaagggg	catacacaca	ccgggcacat	atttctttcc	acccatcccc	aagatcgcaa	122280
gcgcaaaacc	tcgcacagcc	tcacgtttcc	caccagctca	gacatgcacg	ctggcgagct	122340
ttcagcggtc	caccctgtgt	cacactcacg	tgcccccccc	cccgttccc	caagcccgtg	122400
caaagggtta	cgggcaagca	tccctgagtc	cacctgcaca	agcatecttg	cgcgcacgtg	122460
cacgctcata	tgcaactcgat	cttgcaagca	caaactcttg	catatactat	tcttatagtc	122520
gcacactggg	cttgaggtct	gggagtgga	ggaaaagtgg	aatcttggag	ctgtcccagg	122580
ggacagaaat	gctggaggct	gggacactgg	gcgcagggac	gcggctgggg	gcgggggagg	122640
gggtgaccca	gaagctcatc	ttctcctgga	aagtggggag	gggggaacag	gacaagtcca	122700
cggcgttcct	ctaaactacc	gcattccccc	aagaagggat	ttctctagaa	gagtggcgcc	122760
gcgaggacga	tcgaacacag	tccctccgggt	cgcttaagcg	ggggggaggg	gggcgggggtg	122820
gaggggggtta	gaaagccgct	ccgcctcct	agtggctcag	aaagggttaa	gtcggcaagc	122880
cagcaaacga	gggaggagcc	agcgagtgcg	ggaaggagt	gggggtggtg	ggaagagctt	122940
cctcgctgtc	cccactctcc	ctcggctagc	agcctgggca	cacggacaga	cggactgacg	123000
gactctcgag	cggacagcgc	agctagcggg	gcgcggggcg	tgggcgtcga	cggccagccc	123060
cagccttccc	cgccccgtcg	cgccccgccc	cgccccgtcg	gggcccagtg	ctcctcccga	123120
ggccccgagc	ccgggcgggc	cagggtagag	cgccgcggcc	cggccacgca	gcccggggac	123180
tccccggggc	tccccggagc	ccgcgggggtc	cccgccgtgc	atccggcggg	ctcaggggagc	123240
gagtgggagc	gccttcccc	cgctgcccc	tcccccgagc	atcgagacaa	gatgtctccc	123300
gggctcaggc	agctgctgca	aggtaagaa	gccagcgcg	ggagagcgga	gggcatcctg	123360
gggagagaag	cagggcgtcc	cctctttcag	ggattgaggg	tggggcagtt	ggggagggtg	123420
ggtaacctgg	ggaaggggaa	aagctcagcg	ctggggccgc	gcccccgccg	ccagggtgtg	123480
tctcagcagg	agggcacttg	gctgggagcc	cgcgggcgcg	tgcgaggagc	tcgtgaccga	123540
ggtgggacgc	agggggcagg	tggacccggc	ccggagcggg	gagggagggt	caggttccgc	123600
tgtccccgct	ccacctgctc	cgggggagcg	tgaggactcg	ggccggctgg	ggaagcgccg	123660
actcagcaac	tcctcctgcc	cgggtgcctca	gcactttctg	gccacctggg	aagacaggag	123720
atgtgggtag	ggggctgtct	ggggaggtag	gaggcgacaga	gggaaatcca	agtggccctc	123780

tctggtagga	gagatggagg	gcgctagaaa	gaggatagtt	ctactgattg	agtgcacagat	123840
aaggggtgtg	gccagagact	gggggtgggg	tggggagggg	tcagggggag	agggatagga	123900
aggagaactc	aaagatggag	aaagtgggtg	gggaagctca	aaggaggagg	gagatggagc	123960
gggggagggg	gagaaggaat	aaaggttaga	tgggaaaagc	gtggagggaa	gtgggaccca	124020
ggtgaagacc	aaggaagagg	gaaggagagg	aaagaccaga	tcagggggag	gatgggaaga	124080
agactatgga	cagggaccca	gaatcctggg	atggaggtag	cgggaaagag	aatcaggact	124140
gggaccctgg	ggactggaat	ggaaaaggag	aatggaaaag	tcagaaacca	gagaaggatg	124200
gggatgggtg	ctagagaagg	ggtatcagga	accggcgaag	agggttggag	acaggggaacc	124260
atggatggga	gaggggctgg	agaggagggg	agaggaggag	gaagagaaag	gctgagagag	124320
agggactggg	gattgggggt	gctgcccagg	gatgagacaa	agaggcttct	ggtaaccact	124380
tccacgtggg	aagccctcca	ttcccaaagc	gcctgcctgc	cacatttctt	ctctcaggga	124440
gtggctgggtg	ggccagatgg	gggggtgctt	gagctcaggg	ccctgggggt	ggctgtgagg	124500
gacagagggg	gaggactttg	gaaggggagt	gacagcctcc	gaggggtggc	aaacaggctg	124560
gctcctgtgc	tgccatttat	ttatccggcc	cggacgttgg	attctgcagc	cgctgccgcc	124620
accacgggtg	ctgcttattt	tgggggtgta	cattctggca	gagtgagaag	ctgtttgcag	124680
cagctctaaa	cctccgtcac	ccgcgtcagt	gcctccccag	gccccctgct	cactggcatc	124740
accaccacct	ccatcccact	cctcagctcc	cacctcctca	gccccctgcc	cctcagcatc	124800
tgcccgcagg	ccccagccct	tccctgaagc	agcccgttgg	gtgtggagcc	cttgcttctc	124860
gtctgggacc	ctgtgcccct	ccttccagag	cgagaggcct	ctgctgcctt	tccagggagc	124920
atcctttcct	gggaccactc	tgcaccagcg	actctgcctt	gtgggtgggt	agcctggatc	124980
ctgcccccta	ctttgggtcc	agttttcttc	tcctcaagtt	ccttcttcta	caggggcctc	125040
cggcccaaag	agtggcctgt	gggctgagaa	ccttgtttct	gagccttggg	actccaaggt	125100
ttgatagcca	gagtcctgga	cagtggctcc	tcagtgaaca	gatacttttg	gctctggaca	125160
cttcagcctt	ccgggatcaa	taccatgttc	tggctctctc	tggtccctcc	ccctggtcag	125220
ttctggccat	atattctgga	caggggtcat	ctcttcttga	ctccacatg	taatcactac	125280
tctagaacaa	ccgcaactgg	aagcctagga	ggtgaaagtt	gcagagagag	ctggagtccc	125340
ttccttgcc	tgaccctgaa	tagccaaaca	gactcagcat	tgtggctggc	ccagccctag	125400
gcacctgggt	gcaattttct	tcctgtcttt	acctcaaggg	cagtgtctca	cacattcagg	125460
cgtggtttct	gcggaggatg	tggccacctc	ttaaagaaag	atcagagtgt	ctctctgaca	125520
tgggcttgat	gtccctcttt	tccaatctgg	gttccacctt	gtactagctg	gtactctcga	125580
ggccactgtg	tcatgtttct	ggggctccct	tccttcatct	gcaaattggg	ggccacaata	125640
ttgacctcca	ggggattatg	tgtgttgtgt	tcaatgtata	aagaagttaa	cctgtacaaa	125700
tgcagtgcct	aggacaaaat	agggtgcttct	tggtttccct	ctaccctgct	gtactctccc	125760
ctgcagctct	agccatcccc	tgtgtacttt	agaggagggg	gtgagcagag	aggggtggggg	125820
aggctgctac	aaagggtctt	cctctgtcca	tgaagttagt	gagggatgaa	atgaaggctt	125880
ctgagaaaag	caatgaaggc	gagctgtaga	gacctgggtca	ggaggcctgg	ggtgctcaga	125940
aactcacact	tccccctccc	agccctcaat	ggtgttacct	atgatgtgag	gggtcggctc	126000
taggtggcca	ccgaggtatc	ccccctttcca	gctctgatac	tctgtgcate	ttgccccagt	126060
ctccaccggg	aattcacaaa	atgaaggcca	ggagtggagc	cgtgggtcctc	gggagagaca	126120
ggaggcctgg	gcctggaggg	aaggagtggg	ggtgctgagg	aggagtgaga	acaggggggtg	126180
gggaagggac	gtggcaagaa	agaaaagggc	acacactggg	cagggcaggg	actgagggcg	126240
ggggagagag	ggaaaggcac	agctctctag	tcccccaacc	ccccagtccc	accacctctg	126300
ccctggagtg	ctcgtctccag	ccccagcagg	cctggggcag	tgaagcccag	agccccctcc	126360
cctccccctc	tccttgccctc	cagtggagagc	cgctgcgtga	attatggatg	agctccttgg	126420
gttacagctg	ctttgcacgg	cagtggcaag	ggccagaaat	ggcaacagag	tactgttat	126480
gcagcagctg	ttatggagga	gccccagca	ccgggtcgct	cttcagagag	cctgcaggga	126540
ccactatcat	gggctggggg	aggtagagccc	tgggtggggg	agacatggga	acaagatgga	126600
aggagagtgg	ggaaagagaa	gagaagtagt	ctaattgtggg	caggtgggga	gcaggagagt	126660
ctagggagag	aaagaggagt	aggcacccct	gccagctcct	gcagagttaa	ccctcaaggc	126720
cgggaaggaa	cttgatgcca	ggggaatggg	ccttgccctc	gagattgcac	atccttccct	126780
ctgtctctcc	tggggcagcg	gtcagtcagg	aggctggggg	aaagctctgt	aatcctccag	126840
gggctagcgg	ccatcagggc	tcacactctg	gtgagcttgt	ggataagggg	taggattaag	126900
ggatcagaga	aggatttggc	ttcttttggg	gtcaagtcct	tagggaagtg	gagatcagag	126960
ggtgactctg	acaggaaggg	aagtgccctg	gctgggcctc	aagagacttt	tctggccctt	127020
tccctgccaa	cactttgtctg	tgtgaccttg	ggtaagtgcg	ttgctctctc	tgagctccag	127080
tcatcacctc	agtagaactg	atgcttgaac	cagaggaatc	gaggggacct	ttgcggcttt	127140
gaaatctcca	gttctaagcc	ccaaacctca	acctcatga	aacctactca	gggtccccac	127200
tgtgcttcca	cactccacct	ctgcctgggt	cagatgaggg	gtaagagaca	ttgctcctcc	127260
accccacgtg	ggtctaagaa	actcgggagg	agaaagtaat	cgtgaaacgc	cgacaggggg	127320
aggggtgaga	agggccgaga	aacgcggagg	tgggtgtgaac	gaatggaaca	gcagccgctg	127380
tgtcactgag	tattacatca	cacccagcct	acacacgcac	ggggcccgcc	gctcacacac	127440
acgcggagga	cagccagcac	gcacccagcg	agcaccgacg	cagcgccagg	agggggccggg	127500
gacactcacg	gtggggccca	aaagcgagga	gcagcacact	gggagtgtgg	atcttccacc	127560
ccgcacctgt	gtgctccccc	ctctggagga	ggaacaccag	ggcagctggg	atgccagcgc	127620

cacactcggg	gcctgtcagt	cccattgcgtg	cacacctggc	tgagcagcac	tgcatattggt	127680
gagcacctgg	ctcacgccac	tacccaaaat	caacagatata	tacacacatt	cacgcacacg	127740
gcaacctcag	gagcgtgaca	caacacacac	aaaaccacca	ctaagcaagt	gcaatttgca	127800
gccttgagga	ccccacactc	aaaatcacca	accctcagt	ctctcccagg	gtctctgaac	127860
cccaaggagc	cccaggatgt	cagagtgcag	aaacaagtct	tctctcccctc	tgcttcaaaa	127920
agcctaggac	gttgcttgaa	gcagaagggtg	ttcagtcact	gtgtgcccag	ggaatgactg	127980
cctggctttg	gggggtgcagg	ctcccttttt	ccccaggcaa	aactgccaga	agaaaatccc	128040
aggagtcacc	tggaaatcat	aagaaagtgt	agaggccaag	ctagtccgg	cctagaactt	128100
tatcagctat	agtgcaggca	aaggccaggg	atgatgggag	gccctgcacc	cctattaaaa	128160
tatgagtaca	gacacctgca	ctccactctc	tagccccag	gctctctggg	cctgcttttc	128220
catcagtatc	ataataagga	tggatcatat	ccaaccttca	aaagttaactt	tgggggaaaa	128280
aaaaaaaaaa	gctttggctg	gatgcggtag	cttatgcctg	aaatcccaac	actttgggag	128340
gccaagggtg	gaggattgtt	tgaggccagg	agtttgagac	cagactgagc	aacatagcaa	128400
gaccccatgc	ctacaatttt	tttttttttt	tttttttttt	tttgatacag	agtctcgctg	128460
tgtaaccacg	gctggagtgc	agtgggtgcga	tctcggtcca	ctgcaagctc	cgctcccagg	128520
gttcacacca	ttatcgtgtc	tcagcctccc	aagtagtggtg	gactacaggc	gcccgccacc	128580
atgcccggt	aaattttttt	tttgtatttt	tagtagagac	ggggtttcac	cgtgttagcc	128640
aggatggtct	cgatctcctg	acctcgtgat	ccaccgcct	cagactccca	aagtgtctggg	128700
attacaggcg	tgagccaccg	cgcccgccca	aaaattttta	aaaaattagc	tgggtgcagt	128760
ggcacggggc	tgtggtccca	gctcctcagg	aagctgaggc	aggaggattg	cttgagccca	128820
agtgatccaa	gctgcaataa	gctgtgatcg	taccactgca	ctccagcctg	ggcgatggag	128880
caagaccctg	tctccaaaag	aaaaaaagaa	agaagttttt	aagtaactgc	gaatgaggag	128940
agcctggggg	gtaaaatgca	gattcccagg	ctgtccccc	aggaattctg	catagttcct	129000
aggactggct	ggtggcctca	cttagagacc	cgacccttaa	ggccctccc	ggcacaaga	129060
ggctctgact	ctgcaagggc	gaaaagtaca	ggaaagtaag	ggcactgggc	accagtgggc	129120
tggcaagacc	agaccccgag	gtgagtccat	ttcacacggg	cctcagatct	ccaaagggtc	129180
ccaagttact	tccagtcatt	ctccaatggg	gtgactttgc	ccccagggg	acatttgcca	129240
atgtctggag	acattttggt	tgtaacaact	ggaggcaggg	tgctgctggc	atctagtggg	129300
tagaagacag	agatgctgct	aaatgcctta	tatagggtg	ccccacaac	gaggaactat	129360
ccggcccaac	tgtcaatact	gaggcagaga	aacctgacg	ttagtctttt	gacattaatc	129420
tctagacaag	gtcaaacatg	caatagtga	aacaggaatg	aagagatgat	cattcttcaa	129480
ccaatttgca	gtgctttcta	caatggcctt	ttggcattat	tttttaatat	atgagaagcc	129540
tcagaaagtg	gaagtggcca	ggccacttga	ggctataacg	ttgtcccctg	agccccaga	129600
catgggagca	ccagggtctc	aggcctttat	ttttattttc	tattttttcc	cctgaaacag	129660
ggtcttggtg	tgttgcccag	gctggagtgc	aagggtgtga	tcgttgctca	ctacagcctc	129720
aaactcctg	gttcaagcga	tctcctgccc	tcagcctccc	aagtagttgg	gactacaggc	129780
acatgccacc	atgectggct	aatttttttt	tttttcttgt	aaagacaggg	atctccctta	129840
tgttgcccag	gatagtctca	aactcctggc	ctcaagcaat	cctcctgect	tggcctccca	129900
aagtgtctgg	attacagggtg	tgagccacca	tattcagccg	ggtctaggcc	ttttaccaag	129960
ttggggggct	ggcccccagc	tggcactcct	gccctggaag	cccacctagt	aagttctgct	130020
tccccctccc	acagctcccc	cctcggcctg	cctcctgctg	atgctcctgg	ccctgccccct	130080
ggcgggcccc	agctgcccc	tgtctgcaac	ctgtactaca	tccccgcccc	ccgtgagctg	130140
ccaggccaac	aacttctcct	ctgtgcccgt	gtccctgcca	cccagcactc	accagactctt	130200
cctgccaac	aacctcatcc	gcacgctgcg	gccaggcacc	tttgggtcca	acctgctcac	130260
cctgtggctc	ttctccaaca	acctctccac	catctacccg	ggcactttcc	gccacttgca	130320
agccctggag	gagctggacc	tcggtgacaa	ccggcacctg	cgctcgctgg	agcccagacac	130380
cttccagggc	ctggagcggc	tgagtcgct	gcatttgta	cgctgccagc	tcagcagcct	130440
gccccgcaac	atcttccgag	gcctgggtcag	cctgcagtac	ctctacctcc	aggagaacag	130500
cctgtctcac	ctacagggtga	gcctgcccgt	ccccaccct	cagccccctt	ctggtttccct	130560
ctctctgtgg	gccccctctg	tccccgaccc	tgccgtgctg	ccctcctctc	tccccaggcc	130620
accttctctg	cctcagcatc	tccattttct	tctgtctatg	tctcttttct	ctcttacatt	130680
ctccaggggc	tttacttttt	cccttctgct	tctctacctg	tttaggtccc	ttgctgttcc	130740
tctctctctc	tctccctcta	actccacaac	cttcacctct	ctgectctgc	ctgtctgtct	130800
gtctatccct	ttccatccat	cactgcctct	ctcactaact	tgcctcccc	atctgtcttc	130860
tgctcttct	gtctgtctcc	cttcacacac	ccactccgca	tacaccccca	tgtctgtctg	130920
cgtgtgtgta	tctgtctctt	tctgtgatct	cacgtgtttg	ccttcagggc	actctgcctt	130980
ccccagggtg	ccccgcccc	aaggcctttg	cagctgtttt	tctcaccacc	cctcaagtct	131040
gcccacatca	cgggtgaagta	gagagagaag	gcagagccac	agccactggc	atccacagaa	131100
aagttgcgct	tctctccaat	tcactgggca	atgggacggg	agaagccccc	accccttcta	131160
gattcccat	ttccaaacct	gtcatctcaa	tgagggggaa	gaaagaaaag	ggtaaatctc	131220
tgttatgcag	ctggagaatg	gatgctctga	aaatggaagg	aataaccagta	attgttatct	131280
attgttatta	ttattgatct	aattattggt	tattgttgtt	atgctgactg	tttgacagc	131340
aaatccatcc	actccatttc	cccagggaagc	aataacacac	cctccaaacc	acctgagag	131400
aaaatcttcc	cttggtctaca	gagcctccgg	ctggaagggg	gtgaaaatat	ccaaattctg	131460

ccctctccct	acttgaacct	ggaacgtgct	tcctctgcct	catccagggc	tagtgcctaa	131520
ctagttatca	atctgctagt	tggaaaatca	ggtcagtgc	gatgatgcta	atgataataa	131580
caatagccat	aacaacctaa	caaacatact	gagcaccac	tacgagctag	atgctaagaa	131640
tacagtagtg	aacagaacag	accaaaccac	ctgccttcac	agagatacca	ttcccatgag	131700
gagggaaaga	agtaaaatgc	acggtatatt	ggaaaaatat	gtcttatatt	attcttattg	131760
ttgcctaaat	agtgcacgta	atagcagtag	ccgccaccac	ttagtgggta	cacagggtca	131820
gccacagtgc	caagcacttt	ataggtatcc	actctgccat	ttacaagcgt	gtgacatttt	131880
ttttttttac	ctcctcagac	ctcagttttc	tcactctgtac	aatggggtag	caagagcacc	131940
catctcctag	ggatttttgaa	agcattaaat	gcatgaataa	tttgtaaaagc	acttagaata	132000
gtgcttgcca	tacggtaagt	gctatataaa	tgcttgtaaa	aataatattt	taaaaaaaga	132060
aacgagcctt	atttaacatt	ggtttcagtg	aagtggccca	acttggactc	catcctgaag	132120
atgtgggtca	acttcaagga	ttatactaag	gtcatgagtg	agtcccagaa	attgcacctc	132180
acagtttatg	aagtgcactc	agccacctca	tctcatttct	acagcccagt	tgggagatta	132240
ttttcacctc	cttggttaaca	atggagaagc	tgaggctggg	ggccctgaag	accctataga	132300
gatatagtca	cctccaatca	taaactcttt	caacctgttg	cgggtgtgacc	tgaggcttat	132360
gtcttctcac	catcatgttg	agcctcacaa	acacctgttg	atagggacag	ttaggggcac	132420
tagggacatg	gaatgaatgt	tcctgaggcc	acacaccacg	gaagagctgg	cgcttgaacc	132480
tcagtgtctg	gctacaaggg	gacagtactc	tggagtacaa	ttgagcaggc	tcattttttga	132540
aagcacacag	tttggtactca	gcaagaccta	ggttcaaatac	ctggctccta	tatatatgac	132600
tttggaacaaa	ttacttaacc	tctctcagtc	tccatttctc	catctctaaa	atggcaatca	132660
ggatagtact	taataataat	cttttttttt	tgagacgacg	tcccactcta	tcgcccaggc	132720
tggagtgcag	tagtgcgac	tcggctcact	gcaacctctg	cctcccaggc	tcaagtgatt	132780
ttcctgtcac	agcctcctga	gtaactaaga	ttacaggcat	gtgtcactac	accagctat	132840
tttttgtatt	tttagtagag	aagggtttca	ccatgttggc	caggctggtc	ttgaactcct	132900
gacctcaggt	gatccactca	cctcggtctc	ccaaagtgtc	gggattacag	gtgtgagcca	132960
ccatgcccag	ccaataataa	tccttatatta	agaagtgttg	taaggattaa	aatgtaaggc	133020
atttagcaca	aggattaaaa	tgtaaaggcat	ttagcacata	tgggcactat	aataataatt	133080
actactacta	ctactactaa	tactgagatc	aaatactact	acaaattgat	catgcattta	133140
atgctttcaa	aatctcctta	tcaatatata	ttagtatttt	aggaggaatt	tggagtccaga	133200
gggcttgagc	ttgaatcccc	gatctactat	ttcttgactt	atttaacttt	aagcaggttg	133260
ctaaccctct	ctgaacctca	cttactttat	ctgcaaactg	ggaataatga	aaataataacc	133320
ttccaccaag	aatggctgta	aataggaaac	gagtttagtgt	atagaaagcc	catagtccag	133380
gctgggtgtg	tggcccatgt	ctgcaatccc	agcacttcgg	gaggccaagg	tgggtggatc	133440
acttgaagtc	aggagttcga	gaccagcctg	gccaatatgg	tgaaccctg	tctctactaa	133500
aaatacaaaa	attaggcagg	cggggtggga	gggtgtctga	atcccagcca	ctaggggaggc	133560
taaggcagga	gaatcacttg	aacctgggag	ttggagggttg	cagtgaagctg	agatcgtgct	133620
actatactcc	agcctgggtg	acagagcaag	actctgtctc	aaaaaaggaa	aaaaaaaaaa	133680
aaaagcccat	agttcagtg	tgaagaaatc	atgttattat	gaccccatcc	tccattgact	133740
ctcaggccaa	caacagcaat	caggacctga	ggtcagcaaa	ggcttgggca	gaggggacct	133800
cagggtggaca	ttgggggtctt	ctgaaatggg	aagtgtttgt	tctctacgcc	cctggcatga	133860
atggtaccag	gcatcatggg	aaggaagcaa	cttcacacct	ggccttttat	agaggagatg	133920
gaaaacacag	cctctgcctg	tgaactgcct	ggtagggctg	ggctgggaga	tgccacagc	133980
agggtaggaa	acatgggctg	gggtgagatc	cgcaggggtg	agggtgtgacc	caagatggag	134040
ccaggcctgc	cccaaagggg	agctttggag	gaaactccac	cagaggacca	cagcttttca	134100
gaatggggaa	gggccaggca	ctgtgccagg	tgagttcatt	catcaacaga	tatttactga	134160
gtatctacca	catgccaggc	aatgttccag	gtgccaggga	ttcaggagag	aacagaaaca	134220
gtggccctgt	tctcccagag	catattccct	actcaagtgt	agccagatga	taaagacact	134280
tgttttcttt	cttttttttt	tttgagacga	agtctcgctc	tcttgctcag	gctggagtgc	134340
agtggcacga	tctcggtcca	ctgcaacctc	tgctcccag	gttcaagcga	ttctcctgcc	134400
tcagcctccc	aagtagctgg	gattacaggc	atgtgctacc	atgcctggct	aattttttgta	134460
tttttagtag	agacggggtt	tcacatgtc	ggccaggctg	gtcttgaact	cctgaccaca	134520
ggatgatctg	ccaccttggc	ctcccaaagt	gttggtattac	agggtgtgagc	caccgcaccc	134580
gccgacactt	gttttctctt	tcagtcatta	cagtggcctg	catggttttt	gtttgttttg	134640
ttttgttttg	ttttgttttt	tgagacagtc	tcactatgtc	acccagctgg	agtgcagtgg	134700
cgcaactctg	gctcactgca	gcctcacctc	ctgggttcaa	acaattcccc	catcttagcc	134760
tcccagtag	ctggaactac	agacatgtgc	caccatgtcc	agctaatttt	tctattttat	134820
agagacgggg	tttcacatg	ttgccaggc	tggtctcaaa	ctcctgaact	taagcaatcc	134880
acccgcctcg	gcctcccaaa	gtgctgggat	tacaggcatg	agccaccgta	cacagctggc	134940
ctgaatgggt	taaaaatagt	ctttatgtct	aagcagatca	gatctcagtt	tgaattccag	135000
ccacacctct	aatttgctct	atggctttgt	gcaagttatt	taaccactct	gagcctcgat	135060
ggacccatct	gtgaaatggg	gataacctgt	accttggcga	gcaggggttg	tgaggattaa	135120
aggagatact	actgagctca	cagcccaatg	tctgtgtaaa	agtgaagtac	caatgaatgg	135180
tagctatcca	ttaacaccag	ggaggacacc	aactgaagct	cagcaaaata	aaagcacagt	135240
ccaaggtcac	ccagctagta	aggaacatga	cctagaattg	gccaggtct	gtctgactcc	135300

agagtgcagt	tgttcagagg	tctctggagt	tggaagccac	gttcactgc	atattagctg	135360
ttggacccta	ggcgagtcac	ttcacttctc	tgaggctcca	tctcgtaatc	tctgaaatgg	135420
agataataat	agtatccacc	tcataggggt	gtgacaatta	agttactata	taggatctgt	135480
gtagcacaga	gcttggcaca	tggtaaagagc	tcaatcagtt	acctgcttga	caatgctgac	135540
gccgatgatg	acgatgatac	ccatcctaga	ctgatgagct	ctgtaagcgg	gggtgcctgg	135600
cacagagtag	acactcggta	cagctctgtg	gaatgaatga	ggcacatccc	agaactcacc	135660
aattcataaa	aatcagatgc	agatgggagc	ttaaagatca	cctatcctaa	gtcccttggt	135720
tcacagatga	aaagacccag	gcccagagag	gtgcttggag	ctgcgcaagg	tcacacagcc	135780
aagcagctca	tttgattagt	gtcagagcca	agagctggga	gtttggaggg	aggcaaggtt	135840
aagaacagga	tgctgttagg	gaagcaggca	gggatgctgt	gttaagattc	caaatggatg	135900
cagagagctg	tgaaccggcc	agtggggagg	caagggaaat	gtgggtttttg	aaatggaaga	135960
ggatgacttt	agcagaggct	ctcagcccag	agggagggga	gataggaggg	ggagataggg	136020
aggggcgggg	ggagggctag	ggctgtgaaa	gtcaagagct	tattaatgca	tagagaacgg	136080
ttttaacagt	ggagagagga	aggaccggat	ttgaaagcta	cattcaagga	agtggcaacg	136140
ggatttggca	acagcttggg	tggggggagg	aggcaatgga	ccccaaggca	gaggctcaga	136200
gaaggagggc	gcaggacttt	ttgcagagaa	acaaaaggag	aggagaggag	gttagaatca	136260
agaaaattctg	tgggccaana	cctggggctg	tgggtcaaag	gcacctgaat	tccttaggat	136320
ctctggaact	ttgggtctact	cttctgacct	cccagggtcc	cccaaatgt	ggattacccc	136380
tgtcactct	cccccaaccc	ccggcccctt	atcgatcctc	tgaccataca	tctctgggtg	136440
tgtcctactc	ttgctgacac	ttcataaaaa	gaggaacccc	atthaggtgt	tttgagtggc	136500
agggattcca	agcctacccc	ctggatgggc	ctggaagaga	acaagagcac	caggccatgg	136560
tgagtgcagg	tgaggccagg	gaggtgcaag	gagccagctg	gaggcctgag	ccaggtattg	136620
gggttgaggc	acaggggggc	ggagaatggt	ggtgtcagag	gcagccgaga	aggttgaggg	136680
ggacggatct	caatgtggcc	aagaggaggg	ctcttggcac	gctcagttcc	tgtagcgaag	136740
agggcggaag	ccagatggga	gggggcgaga	acaggcagga	gcacaggaag	gtggaggctg	136800
tgggtgtagg	ctgggagtca	atgccctccc	ccaacctgag	gcctccgacc	aggctcctgg	136860
gtggcaggca	tggggaggga	agcgtctccc	caggcagtga	gggagggaga	gccacagtca	136920
gggaacaggc	cccctgggtg	aactggcctg	agcagagtgg	atgctcctgt	tctgagaccc	136980
agacctcctg	gaacctgctg	accacagtga	tgccctgcac	aagaggggag	gacctcaagg	137040
cagtgcaggc	agggagctga	agtcctgctt	ccctctctgg	caagccctta	tctctttgag	137100
ccccagtgt	ctcctctaaa	aaagtgcgt	gggctgatgg	gtgccaaaggc	attagctccc	137160
aagtcagctg	atcatcagaa	tcccctgggtg	agctgggttat	aatgcagagt	ccaggaatcc	137220
ccactggccg	tgggccaacac	acaccgccc	cccccgctg	ttaattctga	accatagtcc	137280
caaggtcctt	tctgcactaa	tgtggcctga	ttaggtgact	ccctagcacc	aggcagggtg	137340
gacagcgct	ctaaggggag	tagtaatgca	atgtggcttc	cttctctctc	tcccctgccc	137400
cctctggggg	tggagctgat	gcccctcacc	cgaataccca	gcctagttagc	agctactttg	137460
ttccccagg	gagctcctct	tttaaagaaa	agggacagga	cccaattgtt	actgagcccc	137520
tattgtcata	gtagccacca	tttattgatg	gttgactatg	cacctgccag	atactgtacc	137580
cttaacagca	tttatcatcc	aaccctcctt	tagcctgctg	aggggggttat	acataataag	137640
gaatattgta	catactgagg	aacctgagac	tccatgaggt	taaaacttgc	ctaaaataac	137700
acagctaggg	aaaaggcaag	ctggattttg	aactagggt	ctaagtgcgt	agcctgtggg	137760
cttcataatg	ggaccaaatc	cctgtgtgct	gggcacgtgt	ccagcacttc	cctcataatga	137820
tctttatgtg	aaccatcctc	tggaaatcctc	agaacaaacc	cagggaagtag	gtatactcat	137880
ccccatttta	cagatgagga	aacaggcaca	gagagatgac	tggcttggcc	aagttaagaa	137940
taatggctaa	caaaacaaaa	caaaaacaaa	aattaaaaaa	aaaaaaagaa	taatggctaa	138000
ctcatggaac	tcatagaact	ccacaaggaa	aggtgttcta	agcaccttca	tacatgctgc	138060
ttcattttaat	ctctacatta	tacagatgag	gaaactgagt	cacagatatc	ctgagtgcct	138120
tgcccacggg	ggcatcagtt	aatgacagat	ccaagatttg	aaatcagaaa	ggctggctcc	138180
ccagtctcca	tacttcacca	aaccagaagt	ctgaaaactc	aaactgtggt	cctgccaatg	138240
gccacactgg	cttcctggg	gaacctgtag	acatggggat	tcccaggctc	caccccaaac	138300
ctcctgaatt	agaaactctg	ccccccgccc	caccccgctc	agagatccgc	aggggatcct	138360
aatacacccg	aaagttagg	aaccactgac	ctcaccaata	ccactttttc	cacagcaaat	138420
aggttagagg	aggcagaatc	caaatccagg	atgctatgaa	tcaaaaggct	aaccctttct	138480
cttctgccac	ggtgcacccc	cttccctccc	ccggccaagg	ccccagcggg	gtctgcaccc	138540
tgccctcaggc	ccattctctt	cttctgtgccc	ccactccacc	ccaccagga	tgacttgttc	138600
gcggacctgg	ccaacctgag	ccacctcttc	ctccacggga	accgcctgcg	gctgctcaca	138660
gagcacgtgt	ttcgcgccct	gggcagcctg	gaccggctgc	tgctgcacgg	gaaccggctg	138720
cagggcgctgc	accgcgcggc	cttccgcggc	ctcagccgcc	tcaccatcct	ctacctgttc	138780
aacaacagcc	tggcctcgct	gcccggcgag	gcgctcgccg	acctgccctc	gctcgagtcc	138840
ctgcggctca	acgctaacc	ctgggcgtgc	gactgccgcg	cgcgccgct	ctgggcctgg	138900
ttccagcgcg	cgcgcggtgc	cagctccgac	gtgacctgcg	ccaccccccc	ggagcgccag	138960
ggccgagacc	tgcgcgcgct	ccgcgaggcc	gacttccagg	cgtgtccgcc	cgcgccaccc	139020
acgcggccgg	gcagccgcgc	ccgcggcaac	agctcctcca	accacctgta	cgggggtggcc	139080
gaggccgggg	cgcccccagc	cgatccctcc	accctctacc	gagatctgcc	tgccgaagac	139140

tcgcgggggc	gccagggcgg	ggacgcgcct	actgaggacg	actactgggg	gggctacggg	139200
ggtgaggacc	agcgagggga	gcagatgtgc	cccggcgctg	cctgccaggc	gcccccgac	139260
tcccagggcc	ctgcgctctc	ggccgggctc	cccagccctc	tgctttgcct	cctgctectg	139320
gtgccccacc	acctctgact	gcggtgctga	gatcgaagag	gccagtgtcc	gatccccgct	139380
tcccgctccac	ccggggctgc	ggctccggcc	ccagtgcgcc	caccttccct	ggccttgctg	139440
cctccctttc	ccctcccagc	tcctctcctc	cccggggagc	aggccgcctc	tccttgccctg	139500
ccccctgggc	tgtcctgact	tgtggcagcc	ccaagagggc	gtgtgtgggtg	gctcagccct	139560
gccctcccca	gttctggcca	ttaactcttc	cccattccaa	ggctgggggtg	gggcccccca	139620
ggcagccgct	gacccgcact	cctaagggcc	cacagcggac	accagagggg	cttttgctg	139680
cagagcgctc	cccaccagca	gagcctttgg	aagctccccc	agggagcccc	accaggacc	139740
ctttggggga	tgccctcagtc	agggccaggc	tgaccctgac	ccctgcttac	cctagtcccc	139800
tcaacctcct	gacactggag	gaatactttt	ctcctaagtc	taccttgga	acttttttagg	139860
gcacctggag	agaactttcc	tctccactgt	ggcccttgcg	tggtgaagat	caaaagaagt	139920
tgtttgggaa	aaaaaattta	ttaaaaaatt	ctattatttt	atctactgta	agatttggtg	139980
acttgggacc	ccgaaagcgg	gatgaggtct	cagaatgtaa	ggattgcagg	gccaggaggg	140040
ttggagaagg	ggagccgtcc	cccgccatca	aagagcttcc	tggtggctgg	aggtgggtgtg	140100
cgctcccccg	ccatgaggag	gagcttattc	cctgcattct	aggtgaggcg	cagtgtggca	140160
gccaagagtg	gggtgctggg	gcacctcttc	tcttcatattg	tccaggggaa	gagctgcagc	140220
caaccttgag	tggtctggcg	cctgaggaac	taagcctggg	gaagacctgc	tgtctgggta	140280
acagccctct	tccagaccct	gttccttcag	gaaacaagag	cagttctcct	gcaaggagga	140340
gtcacatata	cactcctggt	cacagacagc	cccaacatgg	ctttgggtaa	atgtgaacaa	140400
ggcactgctc	cctcagggaa	acacagcccc	atgccagagc	aaacacctta	gcaaacagag	140460
accaaggctg	ggtttcccg	tacacttgcc	tcttggtcta	agtgccttg	tgcagtcac	140520
agcgtaacata	cctgcacaca	gcaacctgtg	gggtatgtgg	tctctctctc	agctcctgtg	140580
aggtagaagc	catcagggat	gaaccaggtc	agagaagcag	gtttccaaac	aggctagaag	140640
agggaccgag	gaactcgggt	gatcagaggg	acaggaatcc	caaattggga	tgcattactg	140700
gcttgaggta	caatcagaac	cttcatcttt	ctggtgtgtg	gaagagaggc	tggggactgg	140760
gaagagctca	ggctaagaag	gacttgggtt	gggatttagg	ggtgagtctc	atcagactga	140820
gcacttggag	agaagtttgg	tagtttgaat	ttggagctaa	gaatctagct	tgggcagggt	140880
gtggctcgct	gcacctgtaa	tcccagctaa	ttgggaggct	gacgtgggag	gatcactga	140940
ggccaagaat	ttgagactag	cctggacaac	atatcgagac	tgagtctctt	aaaaatgttt	141000
ttttaagaat	ctagtttgga	gtggggtgtg	atgtctcaac	gtctgtaatc	ccagcactct	141060
gggaggctga	ggtggacaga	tcacttgagg	tcaggagttc	aagaccagcc	tgcccaacat	141120
ggcagaaacc	ccgtctctac	taaaaattca	aaaaaattag	ccaggcgtga	cggcgggtgc	141180
ctatagtccc	aggtaactcag	gaggctgagg	cacaagaatc	actccagcct	gggtgacaga	141240
gactctgtct	aaaaaaaaaa	aaaatctagc	ttgggagggtg	ggaatagaaa	gatagagggg	141300
gcctagatgc	tagggcttga	ggaagcaggc	tgaggttctg	tgattctggc	tagggaggtc	141360
aaatgatctt	gagaagaaga	gaagaaagga	gaagaaatca	gcattctaagc	ctgaggcagg	141420
tagactccgg	ttaagggtgt	gggggtgggt	gggggagagt	gagagcagct	ggtcagaaac	141480
ccaggagct	cggagtctgg	ggtcttgca	gggcttggtg	caggctggct	gtgaggagg	141540
taatgggttg	gattggaggg	acagccagac	aagagctctg	gtggaggagg	ggctgctggg	141600
gcctgggcag	ggggaggggg	gctgctggta	aattagaggc	aggctgtcca	ggtcatagaa	141660
ttatcattgt	gaaatattca	tgggccatcg	gtccagatgc	tatttcagaa	cagtgaagac	141720
aagaggagt	tgtaggcctc	aggaagaagc	ctggaacaaa	gccactctcc	acccaacccc	141780
acccctccca	ccaccagccc	agacagaccc	acggacgccc	atcacgtgca	caccacact	141840
cccagactct	cacacacact	cgcaccaagc	agagccatgt	agcacgtgca	agcacaccaa	141900
ccaccacagg	gtcccacaaa	caggcagggtg	tcccctaaat	tctgacatgc	acactgacat	141960
gcacacccac	tcaatcagga	cccagcagag	atcacctcca	gcgatctcac	atgcccagac	142020
ccccaaactc	tccaaacaac	ccagattcac	caccttgacc	cacacaccct	gagataggag	142080
ggatgttcaa	ggccatccag	cccaacccc	accaatgctc	tgatggggaa	actgaggcca	142140
tagaaaggaa	gggatttgct	tgagattcct	ctatcccctg	aaaaaagcaa	aattcattca	142200
cctcccacat	tctgagtgtg	ccccattct	gcattttcgt	ctgccagaca	cccagcctag	142260
ttgtaattaa	ctcctccctt	tctctaattt	cctgcatcta	ttcagttacc	cagtccccc	142320
cccagccaca	gtctatccct	tccttcccat	tctccccacc	acctccctgc	tccagctact	142380
cattacctca	tgcttggaat	ataaaaagaa	actgcgataa	cctcctcgct	ggtttcctac	142440
atggaatctc	tcctccctc	ccaccagccc	ataccgtggt	gaccagattc	atctgatcaa	142500
aatttgcata	tgttatgatg	tcactcagga	gctgtgaatg	gcttccta	gcctataggg	142560
taaaggtaaa	acaccttagc	agagcatcaa	agatccctca	gagtctggta	ccaactgctt	142620
ttctagcctt	ttctctcaca	atctcatccc	aaaccttcac	tccagctaga	acgtttgtat	142680
catactggcc	accagttatc	atgtatgtga	aaccaccaa	ccgactttga	gtgccccct	142740
aaaatttctc	agtctctcct	gaagtaggaa	acctcttccc	cctcctcaga	tctcagactc	142800
cagagccctt	tcccaaggcc	aagactgcac	ctctctgacc	atatacaggg	gttcttcaaa	142860
gcagcagaca	gaggctcagg	ctctggctcc	ctccaagcag	acggctgccc	ccgactggcc	142920
accttgggaa	gcacagccag	gtttcagtcg	tctagaacag	agaatgagca	tctaaccgcc	142980

tgaggagagg	actaggacac	cagatgataa	ggtttataag	cccttaagcc	tctaaggttc	143040
ttacacccag	agtagggggg	ggacggttct	cagccctggt	tccctagctg	cgggctccca	143100
atthttcgatc	cctaattccga	gaggaactcc	tctccaatga	aatacagact	tgggactctc	143160
aggacactgt	ggaagggaaa	tttcccaaca	gactctgaga	gtccaggagg	ccagggatag	143220
accaggtggc	aggcccaagg	tccagctggg	gtcagggtttc	tatatgaatt	tttaattgctt	143280
ccagatagac	ttgtcagatg	ttctgaaaac	tgagcatctc	ctttcacctc	tgtacatgat	143340
gccctttctcc	aaccccatg	cccctgcagg	agggcaggcc	tgggacagat	attcagtggtc	143400
ctctggagaa	acggttttgg	gacagtagaa	gggtaaatga	cctagttatg	ttcccactag	143460
taagctgtgt	gaccttgggc	aagttactta	acctctctga	acattagagt	tctgtgggtt	143520
tgtttttgtt	ttgtaagctg	gggacaatag	tgccagccta	aatcaatttg	ttgtggggac	143580
tcagtgaat	agcccatggc	aaagtgcct	acatgcttgc	tgttattatt	ctctttcctc	143640
aagttctgcc	tccctcttcc	agcttttctt	ccaaccccaa	agatgtctct	ggctattgct	143700
tcgaaggtag	gaactttggt	tggttctccc	ctttctcttc	aggcccaaac	tccccacctc	143760
aagatccttt	ggcctttgta	gaaacttcag	gtgaggaggt	ggcagagaaa	taagaaagtg	143820
tgcaaggctg	gtggagtga	agaggaggat	agatggcgaa	gccctagcag	aggggagggg	143880
agtgggcagt	ggagagagg					143899

<210> 16

<211> 215980

<212> DNA

<213> Mus sp.

<220>

<221> modified_base

<222> (1001)..(1100)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (2123)..(2222)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (3728)..(3827)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (5168)..(5267)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (7481)..(7580)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (8849)..(8948)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (10375)..(10474)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (12270)..(12369)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base
<222> (13438)..(13537)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15902)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15939)..(16038)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (18223)..(18322)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (20974)..(21073)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (24403)..(24502)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (27574)..(27673)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30892)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30901)..(31000)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (34443)..(34542)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (38205)..(38304)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42373)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42386)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42393)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42461)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (44809)..(44908)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (51380)..(51479)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56740)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56765)..(56864)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (62818)..(62917)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68518)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68534)..(68633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (74552)..(74651)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (81446)..(81545)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (88519)..(88618)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (93791)
<223> a, t, c, g, other or unknown.

<220>
<221> modified_base
<222> (93794)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96565)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96570)..(96573)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96579)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96590)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96596)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96602)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96616)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96629)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96668)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96715)..(96814)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (104447)..(104546)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114521)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114527)..(114626)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (127063)..(127162)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (139133)..(139232)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (151051)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (153242)..(153341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164706)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164708)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164710)..(164809)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (182242)..(182341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192158)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192192)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (198842)..(198941)
<223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (199437)..(199438)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (208276)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215974)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215976)..(215977)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215979)
 <223> a, t, c, g, other or unknown

<400> 16
 ttggggggtat aaaccagaa gtgggattac tgcaccatac aataatcctc taacttcaag 60
 caattttttcc acaatggttg tatcatttta cattcccact ggctacgaga aggggttccca 120
 cttctacaca tcttcaccac cattttctggt tttgtttttg agtaacagct gcctaagtac 180
 tgtgaagtgg tatcttatct cagtgttgat ttgcatttct ctgacatta atgtgggaag 240
 gcatcgtttc atatgtttat tggctgtttg tgtatcatct tctttggcga tgttgattca 300
 agttatttgc ttgttttttt aattggagtt ttaaaaaatt gttgttgagt tgtgggagtt 360
 cttcattagc tctgcatatt aataccctga tgaaaatgat taacaagtat ttgcttccat 420
 tttgggggct tccattctgg gctgttttta ttcttttgat actcttttga ttctcaacag 480
 tttaatctga ctaaaattca gtttatttct tcttttaatg gccatgctat tgacacatcc 540
 cgtaatcact gccaaatcca gtcatgaaga gtttctttca agagatttat agtttttagct 600
 ctttaagttt gtcatgtctg tttcacttaa ttttgtatag tgtacaaaag tctaacttca 660
 ttcttttcta tatggcttgc tactagtata cgaagagcta aatttctctt tccttgagtc 720
 tcaacctctg atgtgttagc atttcttcag aggaaaacat ggtgggaagt tccttaaaca 780
 taggatgctc catggaggtg aaatagttca tcctacaggg aagcttggtta aacacaggaa 840
 gtacatactc agcagctcta gtaagtgagt gaaactgact ggaggcacta ggtccctcct 900
 tccctacgca tatagaagct gtaaggattg ggaagagata ctgtcaggct agctcagctg 960
 ctgcccggaa gaagctcaga cccactggcc tggctccaag nnnnnnnnnn nnnnnnnnnn 1020
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
 nnnnnnnnnn nnnnnnnnnn atcactcttt actcaggcca cctacacgct gtttatagcc 1140
 tgcttttctc tctttggcta tacttctctg ttatgtctat gcctcccctc tttcttttct 1200
 tttctcttct cttctcatct catctcatct ttcttcaggg gggagcctgg tctagaactc 1260
 acaaagattt gactgtctct gtctccttgc actaatataa aaatctttta caagcatctt 1320
 ttagcaattc ttacaggga attttggaat gttaaactct gattgttagc gggctgaaga 1380
 taacaatagc tctgatgata aattgcttgc caggcaagtg tgaaaatctg agtttgatcc 1440
 aaaaagccgg gtacagaggg caaagagtcc ataatcctag taggggcagg aatcagggat 1500
 ggggtgggtcc ctgggggtttc ctggtttgtc agcgtagccc aattgggaat agccagggtt 1560
 cagtgaacga tgctttctgc aagctgagag aggtccttgt tcaatctctg tgacccaact 1620
 ggaggggagaa gagagccagc tctccagaag tggctcctctc aactttgtgc atgcatgtcc 1680
 atgttcacac aggggaatgga taatgcttaa aaggaagacc ggcagggggg ttgtaatgca 1740
 cctccttttg tgacatgctt tcctcttgtt catgctgtct cagggtgtgg cggcagacc 1800
 aaaaaccagg tgtatgtttg taatccaggt attctctggg cgctagtagg aaatgaaaag 1860
 cgagggtcatc ttctgtataga gttagcaaac tctaagccag cctcggctac atgagacttt 1920
 gtctcaaaac aaaggaaaaa tcaaggagga cggctcccga gcaactgtcac ctgaagctga 1980
 cctctggcct ccacatgcat gtgcgcaaac acatgtcctg cacaacaca cagacacccg 2040
 catctgctcc ccgacaaaag aacctgaaac cagtatactt tgagaatttc ccattcatag 2100
 ttaccattgt gtgttccttg tgnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
 nnggtggtgc ctttctctta cccagctag aagggtgga ggcaggggtg atggggcact 2280

ttgaactccc	acctaggcaa	aaaaccagtg	gatctctggg	ccagtggtgt	gtttgcaagg	2340
gaataaggta	gagagccgcg	gaggaagaca	ttgggggttc	tatgagtagt	tgaaggggtg	2400
cacacaccac	acacacacat	tttttttgtt	ttaaatttac	aaacattaaa	ataggctgta	2460
atgtgggtca	gtgggtagaa	aaacctgctg	tctaagcctg	gtacgagttc	aatccctgac	2520
aagctggaag	gagacaacca	accacaactg	ctagcagcca	gaagcactgc	ttgctaacac	2580
tcaagagagc	ctggagtggg	agacactgga	tccccagcag	gcaagcctgc	aagaagatgt	2640
gccttgcccta	gacaacggca	gaacaaacat	caaggctggc	agagctgtcc	aggactgttc	2700
atattaatca	tgtatagata	agaggggaatg	gcacagacag	aacaattcaa	cacacggggt	2760
atgaaaggaa	aggaacaagg	cacacaaaagg	acaaagaacc	tagcatacaa	gaaagcctaa	2820
gcagagagt	gcacttccca	gaagggaatc	ataaaataga	ctgaattcat	taaaacaaga	2880
gccaagata	aacggctcaa	aaaactcacg	gaaaacaggt	caaaataacg	tcacccatct	2940
gacagttgat	actgtcaact	taaccgtatc	tagaactcca	gcaggcacat	ctccaggcat	3000
gcccctgaag	gggtccttgg	actaggttaa	ctgacgtggg	agtgcacca	tctatggacc	3060
gaagcctcag	acagaataaa	aaggagccag	tgagctgagc	gtcagtgctc	attgcttctg	3120
gcttcctgtc	tgtggctgca	gcgagacacg	gtgcttctctg	ctttagctgc	catgacagac	3180
cacaccctca	aaccgtgaac	caaaataaacc	tcctctctac	attgctttta	ccaggcattt	3240
ggtcacacca	atgagaaaagg	ttaactaata	cagcactcaa	tacttaaaaa	caggcaacc	3300
aaacttggtt	gcattgtgtga	gactttgaag	ctcacggggc	agttatgccc	aatgccaggt	3360
ctgctggcta	agggtgagag	tgcacaccta	taatcccagc	tgctgtggaa	tcagcaaaag	3420
cgctacagat	ggaaggcagc	cagggcagct	gagactgact	caaactgata	gaggtgggag	3480
gcatagagaa	aaccagatta	atagagtgtt	ccccactatg	caagaagccc	tgggtttcag	3540
gacgagagaa	ctaagaatac	agaagtctac	tgtgtagaag	cactgctagg	tcacacagaa	3600
acatcactca	agtgtctctg	gatgtctac	ggaggcgctg	tgaagtattg	cttctctatg	3660
atctgtatct	actacagcac	tgtgttttta	tgtgcgctc	ctccactaca	gctcctcacc	3720
acaccaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaat	taatcaaaga	3840
aaacacacac	caccagttag	agaaagttaa	tcaggccgaa	tggcggtttt	cccctgtatc	3900
caggctaccg	tcaggacggc	tcaactgccac	tggcaactct	gcctgaacaa	agcccgcagc	3960
caacgtgggc	ttcaggggct	ctaaacactg	caatcaaagg	ttgtgtgtgg	gggtgggggt	4020
gctgctgcta	ttcaaggatt	cccaaagctt	agatgtattc	atcatactca	caggaaagcg	4080
tgttcaaccc	atcactcatg	agcagtcggt	accgggtgga	cctattccct	gtagaattgg	4140
gacgcatgtt	ctggaaaagt	tgacagaaaa	gttgattcat	taggcaggct	ctttgcccaa	4200
gccctgaggg	taagcaaagc	taactggcag	gagactaggt	ttgccattaa	tctgagacaa	4260
gatgaaccac	ttgcccatcc	tctgacacc	taaatactaa	tgaaagaaca	atggattgag	4320
ctggcattat	taaaaacgat	agaaacagaa	gtatcaatag	tcattgtgtc	tttctcccat	4380
atgtcaaaac	aatgtgtaag	atggcatcga	acacatgcag	aaactgttta	gggaacatgc	4440
tgaaaatatg	aagtaaaatt	aaaattggaa	agaaagacaa	tttgccctaaa	gcagctcaga	4500
gctggagaag	ggaccgaggc	agagataaca	gcaacgtgtg	gacatacgga	tctggggcag	4560
agcagtcacg	gactcagccg	gaaaggggtg	ggcagcctct	gaagggaagt	aaggtaataa	4620
gagccacaag	gtgattggcc	caggagtggg	gccaccttca	cctcctgcct	caaagtctga	4680
aggaatgatc	ctggagtctc	ccatctattg	atatatgaaa	ttcacagtat	gttttagaac	4740
ccactgaatg	atgggtagat	taactaaaag	aaattttaagc	ggggtggtgc	aggtccttta	4800
atcccagcac	ttgggaggga	gaggcaggtg	gatctctgtg	agttcgaggc	cagcctgggt	4860
ccaggacagc	cagagataca	tagagaaacc	ctgactcgaa	aaaacaaaat	taaaagctca	4920
tcaaaacaac	aacaacaaca	aaaaaaaaca	aaaaacaaaa	caaaacaaca	aaacacccta	4980
tagtacctgt	tgggtgagttt	gagtgtgtga	gtgagtgtgt	gttagagaga	ggggcgggga	5040
aagtgtgttc	tggaaatggg	agaaagagaa	tgtgcatgtg	tgtttctggg	atgtagacaa	5100
aactacatgt	cttccatcaa	atgcaatgtt	taattatcta	tgagttgaac	catcttcatt	5160
ctgctaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaac	aaaaataaac	5280
caaaccagta	aacaaaatcc	tgtaagataa	agcctaagac	aagacacttc	ctggggctgg	5340
ggagttgctt	agaccataag	gagttcataa	tccaggcggtg	agagcccagag	ttcagggtccc	5400
tgggcttcca	agtcaggagg	agaccaagga	atcaacaagt	ctcgactttg	gtctctagtc	5460
ccatgcacac	acacgcgtgt	aaatacgtag	atgttcaactc	acacacagaa	gactgcacct	5520
ggctctctca	catctcagcc	aacatataaa	gcctgcatta	tcagaacatt	ctaggttcta	5580
gtttcagtc	actcttacac	agaatggcca	tcataactccg	tctacaactt	ctctctgact	5640
acccagctgt	cattgcttca	gtattaaaca	aaccagagat	aaccagctgc	gtgatctctc	5700
cctgatggcc	cagtcattgt	cttactgaga	ctactaagtc	acaaggtagc	actctggatc	5760
caaaaagcaa	tatccaattg	agagttacaa	cctataagga	ggagtttacc	ttcattatag	5820
ggcactggat	tcccaatctt	taatccaacg	tcttcagcag	atttcataac	ttccaagtcc	5880
atcaaaacaa	ctactttcct	acaaagacag	acacaagtta	gaatttaagaa	ctctgcagcc	5940
tttcagatga	gttactaaga	agcttacttt	agtagttgtc	tggctaaaac	tgtatccttt	6000
accaaccttt	tctcattctg	gactaacttg	agaagtatta	attcctaagt	aaatacttca	6060
cttattcttt	ccccacatct	ccaatgtttt	tgtctttaat	ttattatagg	gcaattcatt	6120

tcctatctag	ttccctgatt	aaaacagtag	accttgctgc	atgccattat	cctcatggag	6180
gcactgatac	aatttagatt	attaaatata	aaaccctaaa	acacaaaaag	atgatttttt	6240
tttaaaacaa	gatttttaaaa	aaagcatgtg	ctacgcttcc	ttctgccact	aagcctacac	6300
atggtcctct	gactgaattt	ttccctcat	tctgcttcat	ctaataatgtg	cttttcaaac	6360
ctggaattga	accagggact	tattcatgct	aggcaaagtc	tctaccatag	agctataccc	6420
ctccaactcc	catctcaaat	atcattttcca	aagacatttt	cttggtctct	tatttagatc	6480
aggtttcttt	gtcctcctgc	agctatgact	tcattccttc	agaacactcg	tcttagcttt	6540
aagttctgta	ttaatttagt	attgttttca	ttctctctgc	tagaatgcac	tttcaataaa	6600
ggcaggtagc	cagccacagt	gcttaattaa	gcaacagccc	aacgatgtca	ttcactacat	6660
actgggacaa	gatgcctaac	atcatctgca	gataaagacg	aactactggt	gtcaggagac	6720
agctaagggg	tccagggtct	gggcacgctg	agtgtgagca	ctggagtccg	ggtgcccaga	6780
aacgcacata	aatgcaatat	ggatgtggca	atctacctct	aattccttct	ttaagacagt	6840
ggctctccag	agcaagctgg	ctagcaagac	aagccatatc	agtgtgctct	gggcttgacc	6900
aagaccctgc	ctccagggtgt	aactcccaag	caaaaggatg	atggctcaca	aatctcaggc	6960
tatcatgttc	atgtacaaaa	tgtcaaccgg	catacacaca	tgcacacaca	tgaaaactgg	7020
gagaaaaata	gaagaattgc	aaccaaaaaa	tgtaatttga	ggacacataa	tgcaggcgga	7080
ggagtggggg	gatgacagaa	ggtgaactga	gtggaccgag	ggaaagctgt	gctacggcca	7140
atgagaagaa	gggtggggca	gtctgagcaa	gggttcagca	atcaccacgc	tttactgtct	7200
gcacagcctg	gctgtagaat	gctgggcttt	atcacacaga	attattcagt	atgtgctatc	7260
tttacagtaa	agttattcta	tcaggctatg	ctacttcaat	agaacaagcc	tgaaaaagtg	7320
gtctgctgct	gagaacctga	caaagatgac	ctgttagaac	tgtctgccaa	gtgtggaatt	7380
ccagcactgg	ggaccaggag	ctcgagggtc	acccagatg	caggaggtta	gaggccagtc	7440
ttggcaacat	aacatcatgc	ttcagaaatt	aaaaacaaaa	nnnnnnnnnn	nnnnnnnnnn	7500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7560
nnnnnnnnnn	nnnnnnnnnn	catgagatag	ttaataaaact	gaagaaagcc	atacaaggag	7620
taaagtagat	agttgcaagc	atgaagaaag	acaaaccact	tgagcttttc	ttttgtcgta	7680
aggaggaaac	cagacaggtc	cagagagatg	gctcagagat	taagagcact	gactgctctt	7740
ccgaagggtcc	tgagttcaaa	tcccagtaac	cacatgggtg	ctcacaacca	tctgtacagc	7800
tacggtgtac	tcatatacat	taaataaata	aataaataaa	taaataaata	aatcttataa	7860
gaaaaaaaaa	aaaaacctaa	ccaatcagcc	aggcgatggt	gacacatgtc	tttaatccca	7920
gcacttggga	ggcagagaca	ggtggatttc	tgagttcgag	gccagcctgt	tcttcagagt	7980
gagttccagg	acagccaggg	tgatacagag	aaaccctgtc	tcaaaaaaca	aacaaacaaa	8040
caaacaaaca	aacaaaaaag	gaggaagcca	gacaggatgc	actttatacg	tgaatggaat	8100
tgacaaaaga	caagttctat	aagtgttagg	gaaaggggga	ggacaacggg	ggttcatgtc	8160
tgtggtggaa	cacgtattag	aaggctctgg	gtatcctgtt	tccgacaaac	aggcactccc	8220
aatcacacag	gccactggat	gtctcaggga	gagaaagatg	tgatagattg	actttttaac	8280
aatcacagac	tgtgtggaaa	atatttgtaa	ggttgtcatt	gtcaccagg	atagagctga	8340
tggttattca	aacgaggatg	ggacaacaga	aatgggagag	agggatgtga	gaaccatttt	8400
gaaccagggt	gatttactgc	gcacgtgtat	agggtctaca	gggagtggga	tatgtagagg	8460
aggcctatgt	tcctaaacttt	ggtaatgagc	ttattacagt	tactatgcac	agcctggaag	8520
atactggaaa	aggtgcaggc	taggctagaa	aggtaactaac	tgagggtttg	acagcccctt	8580
ggatgtcagg	atgcagcaag	cctacctctg	tatgtagtca	atcccttctc	aggctatggg	8640
tcttcagat	catccgtctc	tgtatccatt	attcccagtc	catcctctga	gtggctccct	8700
cttatccagt	ttaacaaaaat	gctgactgca	agctcccaag	cccagggtc	tgtctccttt	8760
actccttggt	attgtacttt	accctgtttg	cttgggatag	agtgtgccct	ttataaacat	8820
ttgtgaaagg	gggaatgaag	aagaataann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8940
nnnnnnnnnn	agagctcaat	ggttaggagc	actggatgct	cttccaaagg	tctccagttc	9000
aattcccagc	atcaccatgg	cagctcacia	ctgtctgcaa	ttccagttcc	aggggattca	9060
acactcagaa	acataagtgt	aggcaatcta	cgtaacataa	aaataaataa	atgagctgga	9120
aaagaaaaca	tgttttcaaaa	tatacaagta	atggggctgg	aggagatgtc	tcaatgggta	9180
agatcattgg	ctgctctttt	ggagggttctg	ggttcaattc	ccaccaccca	catgacagct	9240
cacaactgtc	tgtaactttg	gtcctgtggg	agctgatgcc	ctcttctggt	gtgcagacat	9300
acatgtagac	aaaacacctg	catacataaa	ataagttttt	aaaaaagtta	cacatacacc	9360
cgtgtgtaat	ataacacaca	ctggcttaac	ttcctcagca	ctgactgttc	accatacgga	9420
ttcccatgag	gttttggttg	cattctatca	cggaaaaaaa	aaaaaaataa	ttagaagaaa	9480
gtatatacat	ataaacctct	ccctaaaata	aagttttctt	ttctaaaagt	acatccttat	9540
ttttttatct	tttttttttt	ttaagaaatg	ggaacaacag	ttctgctcac	actgtatttc	9600
tagcatgtaa	catcttgcaa	gtacttaacc	gtattctata	tcagctcaac	acacttacta	9660
ccgaagactc	aagatcacaa	aaaaaaaaaa	aggaccacga	ctggataaatt	aaacgtttct	9720
tttggtgtag	taagcgacct	cttccttaga	agatactaca	gtaatgctga	agaaatgaca	9780
catctactgt	aatctgttct	ctgggattcc	aactgttttc	ctctgctact	cctcccttgg	9840
cggcaatggt	cgtctgcac	cggctgagct	cctcgctgcc	ttgttaaacc	tcttctctga	9900
acttccgacc	tgtagttccc	gctctacagt	gcaagcgagt	ggataaggaa	gcgcatacct	9960

gccgtctttc	aggggtgttga	cgatgaactt	gtggacctgg	cagacacagt	tgctggccag	10020
ctgccctccc	tcgaccaggg	tgttcagctg	cgtggccagc	atgaacgctg	caaaagcaga	10080
gagagagggg	ctcagtcctc	aagcctttcc	ttaccgccaa	agctcatcac	aaggagaacc	10140
attaaataca	gctgtttaaa	actcctccgc	cctgcagaga	ggaaagcagc	atcaatccgc	10200
cccatgtaaa	agtctgaggc	tcttcctaaa	tggtatctgt	ttctcacagt	ctccaaatca	10260
tttttactgt	aattctagtt	tctggggaaa	gacctttctc	ggtcttttagc	cccgtgacta	10320
gagacaacag	gcaaattattc	cagaaaggcc	cccattttct	ttttaaagct	tctannnnnn	10380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnngcacat	cttgtgaagt	gtccacatct	10500
ttcggtccct	cgaatttggg	tttcttctgg	gacgtggtag	catgtgactg	tccatccagt	10560
gcttggagca	gcagaggggt	caggaaactcc	aggctggcat	tagctgcaga	gctggagcag	10620
gtcctggaga	acagaaactt	tggttgagc	attaatgaac	tagaagaatt	tttttgtctt	10680
ctgttaaata	taaataacctc	cattatcttc	tcataaacag	tgttgccttt	ttattttaagt	10740
ttttaaggat	caggcacaga	gactccatgc	cagactacca	ctcaaccact	gagctacacc	10800
cccaacttgc	ctttctgcta	tttttttaaat	tgtatcagtg	gccaccaaac	atggggagag	10860
gtcagggggc	tacgtggagg	aattgtttct	ctcctacca	gtgggcccc	ggttccaat	10920
tcaggtgacc	tggcttggca	gcaagcacct	ttacccttaa	gccatctcat	tggtctcatc	10980
ttttaatggc	cccttcccc	gctctgaggc	aggctctccc	tatatagccc	tggtctggcct	11040
caggctcgca	ggtccaccag	tgagcaccag	gtttctgctt	gtccttacct	ccccagcact	11100
gtggttataa	gcatgtgcca	ctgtgtcaaa	ctcagtcact	aagctttgcc	aagccatagc	11160
ccagcccttg	agtttactgt	ttgtctgtgt	ggtgatttgt	caaaccactt	ttgttccact	11220
gaggtatttt	gtcaagtttg	acaaaattag	ttgagtatgt	aggtcttttt	ttctggaatc	11280
ttctgttata	gcttagtctg	gtcttgaact	catgatcttg	cctcaacctc	acgattattg	11340
aggattattg	ataggacag	gctgtgtgac	catgctcggc	tgtgtgtttt	agcatgcatt	11400
agtcatttga	aaaacgttgg	ctcatgacac	tttacaggtc	ttccatgttt	gatatgtttt	11460
atttaattcca	aagtaattcc	agcaccagag	gctgagacag	gaggatctca	aggtcaacct	11520
agagatgcat	agcaggcggg	gccccactcg	gttaggttaa	tatcatcact	gacttcagga	11580
gaaaagtctt	aagtattggg	gactaaaagc	aggaggatct	gaagttcaag	gtcatcttta	11640
ggaacttagc	agacttgagg	ccagcttggg	cgctgtggga	ccctgttttt	aaaccagaaa	11700
acaaattgaa	aggaaaaaaa	aaaaaagctg	gaggaagtga	atgtgagtgt	tcacatagtc	11760
ctgtttccac	aagaaaaacag	ggttactttt	ggcaacaaat	agggtctttc	tttgaaggct	11820
ggcatttttg	tgacttgtca	ttggagaaat	gatttaatta	agacttttct	actgagtgcc	11880
tctgaagagg	ctctttttaa	tttagtttaa	ttttatctca	ttgttagtgt	ggtgtgcttg	11940
tgcacacaga	aggcagcttt	ctagagtctt	ttcactctct	cctccacagc	tcctggagtc	12000
aaactcaggc	cctggctagg	caagctctta	ggacagtgtt	agctgtagct	tattaagttt	12060
ttaagaattt	ttataagact	ctgtttttct	ttctcaggtc	atgatacagc	aggaaaatac	12120
atccataaag	cccacctgac	aggtcattgt	aagtaaccggc	atgtgtgttt	agcataatga	12180
agatggttca	cttatagtta	attaaacatt	ggattggatg	gaagacatgt	agttttgggt	12240
acttcccaga	aacacaaatg	cacattcttn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12360
nnnnnnnnng	aattcagagc	tgatatgtag	tactaactcc	tactcaatga	atcctttgtt	12420
cttctattcc	ttcattacat	tactgttaat	agtggtaact	atgtaccaaa	gagtcaaata	12480
actcttggac	catccaaggc	agaaggaagg	ctggcaaaaa	tgtatgatga	tctgggatgg	12540
gaatgtactt	cagtttgtac	aggagccct	tggttctatc	catttctggc	aatgcataga	12600
cctgtaggat	ctcagcactg	gtgggggggtg	gggggtgagg	gtgaaggggc	gggaggttaa	12660
aggcagaata	gtcataaatt	caaagtctgg	gtcctggaaa	gaggactaaa	cgattaagag	12720
cttttagctgt	tcttctagag	aacctgggtg	gatccccagc	acatgggtgc	tcacgactgt	12780
ccgaaactct	gattctaggg	ggatctgaaa	accctcttct	gccctctgta	gatacagaac	12840
acacatgggtg	cacatacata	catgcaacct	aaacaaccca	tatacataaa	atattttttt	12900
ttcaaaaaga	cattcaaatt	cttcctcggc	tatatagtgt	ttaccaaaacc	tcaaaaacaa	12960
aacaaaacaa	aacaaaacaa	agaatcatta	atgttttggc	ttcatgtatg	tctgccacc	13020
acggacatgc	ctggtaccca	gggagattaa	aagaagacat	tagctccctt	ggaatggaga	13080
taggtatgat	ctaccacttg	ggtgctggga	acctgggtcc	cctgcaaaaag	cagtaaactt	13140
ttttaacccc	taagctgtct	ctcccaacgc	ctaaagattc	ttgtaacaca	gcatgatgag	13200
cactggcaag	catagcatgg	taatctgact	tcagggcgcc	agatttttag	cttaattgctt	13260
gattattaga	agtaacgtac	tagatttaat	gcctggagct	tcaagcaaca	aaattaactg	13320
agaataaaaa	ataaaaaccc	tgccagccat	gatggtaatc	ccagaacttg	agaggcagag	13380
gcaggtgac	tctgtgtttt	gcaaggccag	ccacaatcta	catagcacgt	tcaggtannn	13440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnncga	ataaatctac	acatgtaaaa	13560
agaaattcaa	agaaacaaat	gccaaataaa	tacacatatt	gtaataaaga	gataattgtc	13620
taaaaaactc	aaggctttaa	atggtaagat	atcatattct	tggtatgaaa	gatctaattg	13680
caaaatatat	caatttaatg	caattatgta	tattcaggag	atctctggtt	ggcttttgaa	13740
cttgatagca	ctcttataat	tcacatagaa	gaaaaaaaac	catgaaaact	gccaaacatt	13800

attagaatac	tccacagatg	gtatttttggc	agcacatata	tcgaagggct	gtgaaagatg	13860
tgtagatcat	ccacgccttg	ctagggagag	ggcgggtgtg	tgtggggggt	atagctgttt	13920
gggaaaataa	cctggttaatt	cctcattagt	taaatcatag	tcagaacctg	gactagcaac	13980
ttctctctaa	aatacattca	ccctcagcat	ctgcattgcc	aggaaaccac	tcctagcagg	14040
atctgtacgt	ggatcaaggt	agtagcatct	gcattttaatt	gacattctcc	taaagtcttt	14100
aaattatctc	tagattactt	atagtagcca	agatgatgca	aattatgtta	cactgtatta	14160
tctggggcgt	aacaagaaaa	tgtctctact	cagggttcatt	cagggtgcagt	acttccccctg	14220
aatacttctg	aatacacgga	tcaagaagcc	acagaaagag	ggctaaccat	atacaagcat	14280
atagtacact	aataaaccatg	tacaaccata	tagtacctata	atattcagtg	cattactcaa	14340
aatgcaaaca	gatggaaaca	atccaacagc	ctgtaagctg	aaaaacaaga	taagcaaaat	14400
gtgctggggc	tagaggccca	ggtctataat	tccaactaag	gtcgaggcag	gaggatctca	14460
agttcaaggc	cagcctagac	aacttagcaa	gacctgtgtc	caaaacaaaa	agtaaagagg	14520
ctgaggatat	agctcagtat	agagcatctg	cttagcatgt	gcactgacag	ccgtatcaca	14580
gaggaaaaaa	aaaaataaagc	aaaatgtgat	ctgtctgcac	aacaggatat	cacagccccc	14640
taccacaggg	gaacgacaca	gtaacacaac	aaaaacttag	ccctgaaaaat	actatggtaa	14700
ataaagaagt	gtcactgagg	atcaggaaat	gcatgactcc	atttacatta	tatagaaatg	14760
agaagatcag	tgagcctcta	ggactcaaga	gatttgggat	tggcagctaa	agggtagctg	14820
gtttctttat	gggggttaaga	aaacattcta	aacttaactg	tgagaatgac	tactcaacaa	14880
tgtcaagtgt	tcaaaaatca	tacttttttt	tttttttggg	ttttcaagac	agggttttctc	14940
tgtgcagtc	tggaaactcac	tctgtagacc	aggctggcct	cgaattcaga	gattcacctg	15000
cctctgcctc	ccaagtgtctg	ggattacagg	catgcgccac	cattgtccgg	ctcaaaatca	15060
tacttttaaa	aattgcccag	tgactcatga	atacaatcag	aggcgggaga	ggacagtggc	15120
aaactcagga	taccagtgtc	ttttatgtct	gctgcccaac	tatcaatttc	ccatagttaac	15180
cagagaactt	tttggtttgt	ttcatcttat	ttgttgcttt	tggtagaatc	tcaatatagt	15240
aagatacaag	gctggcctca	tactatatag	ctgaggacga	ctttgaactt	ctaatectcc	15300
tgcttccatc	tcccaagtgg	tgggattaca	ggggtgtacc	gctatgccca	gcaagcacia	15360
agccatttga	accacacccc	agccttttca	gagaaacctg	tacaagcctt	agtgccttag	15420
catattaagg	caacaaaaga	cataatgcgt	ggctaccata	gagtgtttgc	ctaccatgtg	15480
tgaggctcta	ggctaaatgt	ccagcactta	tcaaaaaagag	ttaaaaacac	tcagtactca	15540
aggatgacta	tgcatgtctg	tgtacaaagc	cccgacttca	atccccagca	ccgtgcacat	15600
caggcaggct	ctgtagagga	cccagcttaa	ggctactcct	aggtaagtta	gggccttag	15660
atggctacat	tagatgagac	cctttctcat	aaacagaata	aataatttaa	agctcctgat	15720
caaacactat	gccttcccat	cacactcaga	ataaagcact	ctactggccc	tttaaggact	15780
gcccatctgg	aagagaaacc	taagttacat	tccttgcttg	tgtcatatgt	gataacaaac	15840
tacttgaaa	tacgaaaata	cagtcttaag	ccttggtcaga	aagcttcccc	agcaacatga	15900
tntcagagga	cataatgcag	aaagtggaca	aatgcaaann	nnnnnnnnnn	nnnnnnnnnn	15960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16020
nnnnnnnnnn	nnnnnnnnnaa	tcagaggaca	tctttcagga	gttagttctt	tcctccctct	16080
atagctttca	gggatcaaac	tcaagtgtgt	actgagcgct	tatgcccgat	gcgccatcgc	16140
accaggcctg	cttcttttgt	ttttatgggt	ctgaatcaat	tagcaccatt	acaacaatgt	16200
tgacaatcag	caagtacctt	tctctacctg	gctagtaaga	gaagtaagtg	cctttgtgtg	16260
gtgaacgcag	tttctcttgt	gaagtgcagt	gacttgatct	ttgctcacia	cgttttttag	16320
gtccttaagt	tgcttgggtt	ttatgggaaa	ggctcttggg	ttttttgaaa	agatttttact	16380
acaacttgat	ataatcatta	tttttaatcc	tttaaatagt	atgacttatt	ttaacagatt	16440
aatattgaac	tgttctttca	ttcttacata	ataaatcctg	ccttaaaaaa	taatectctt	16500
agcttccctt	ctctattttc	aaatttggtt	tatatttttg	catgattttg	aacattttata	16560
aaagtaggca	gacaacacag	tagaaccaag	tccccatata	gctgtgcaca	tagcttcaga	16620
ttattgcctg	ataataggtc	ctgttttgtc	tctgttttct	cacagggatg	tgttatttgt	16680
tgtgtgtaca	catacatata	tatatgtatg	tatgtatgta	tgtatgtatg	taatgaactc	16740
cctttaacaa	aacaagtact	gggctgggaa	gacagcacag	ttagttatgt	gtttaaccgc	16800
acaagcatga	caaccagagt	tgagatcccc	accaaccgca	taaaaagctg	ggcatagtgg	16860
cattgacctg	tagccctggg	gctggatgaa	agctggggag	gcaggtagat	cggcagagct	16920
tactggcaac	aaatctgccc	agtaggtaag	ctctgggctc	agacatccta	tataggaaaa	16980
agatgaaggg	cgaggcgag	cggcacacac	ccttcgtggg	agtgccttag	aggcaggggc	17040
aggccagtct	ctgtgaccag	cagcctggcc	tacatgtcaa	gttgaggag	agccagagcc	17100
accacctact	gagactgtct	cagaaataag	ttttttaaaa	aattgagatg	aaggagctgg	17160
taagatggct	tagaaggtaa	aggcacttat	cactaagcct	gaagccccga	gtttgacct	17220
ggacccccaca	ctgtagaacc	aactcctcca	agttcttctc	agacctccag	cagagcacia	17280
gtgtatgcag	acacacacac	taagtaagtg	aatgtaaaaa	acatgacgta	gtggcactgg	17340
cctttaaac	cagcattggg	aggaagaagc	gggtggatct	cttgagtttg	agaccagctt	17400
agcctacata	aggaatttca	aggcagccag	ggctacctag	aaagtagctg	tttatgaatg	17460
aatgaataga	aggaaggaag	aaagagagac	agacttaaaa	aatatatgct	ggagagtaac	17520
agaagaggag	accggcttgc	tggtgtcttg	acctctggct	tgtacacata	cacatgtgta	17580
gtgcatacac	ccacatacaa	ttgtactcag	acacacacaa	acatgtactc	attcatatac	17640

tgcacacctc	aacactcaga	aaatgaaaaa	acaggtacca	tttacacctc	cgtgttcggt	17700
ttccaaccac	tcatatgtat	gggttgtaaa	tgcttatatc	tgtatgtgtc	tgtatatattg	17760
tgtatacatt	caaagttgag	tcaggatcca	acgtaaaact	ggatagtagt	gggttgatgg	17820
tctggaaagc	tgctcgcagc	tgtctttttc	tcctcgtacc	ttttccctcg	tttgtttcta	17880
cgacagcagg	tcatttgtct	ctaagtgtta	gtttcccatc	ctctctcttt	tgctgatggg	17940
agcctttag	tagtcacctg	tgttctctgt	aaaatggcct	tgccgtgtta	tttcaatatg	18000
ctatcatcct	catcttgcta	tatttcattc	aatatatgta	tatattacaa	gatagattaa	18060
aattatttta	attttatgct	tatgaatggt	ttgcctaagt	atattgcacc	ttgtgtgtct	18120
agtgtccaca	gaactcagaa	gaaagtgtca	catattctgg	aactggaatt	gcaggtgggt	18180
gtaagccacc	atgtgggacc	tggaaaccaa	atccaggcgc	cannnnnnnn	nnnnnnnnnn	18240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18300
nnnnnnnnnn	nnnnnnnnnn	nncccttggt	caaagatcct	cagtttcgac	tttgattacc	18360
cagacttctc	gtttctctca	tggaaacagt	tccccctgag	atttactagt	ggaagaaagg	18420
cactcaaaaa	gcagggagcc	ctcgtacaaa	tgagacttcc	tagctatata	attaggccga	18480
gatgcacaca	tccacagtca	ttacccttct	tcagagcctt	tgtcatgtca	agtgtatttc	18540
gccccatgtga	actttagaac	tggcttggtg	tgtttcataa	aaagtgtagt	tgggctcttg	18600
attgggattg	tgttaaattt	atagattcat	ttagggagag	ctgacaactt	tacagtatta	18660
aatgttttca	tccgcggaaa	aggttgtcct	gccacttact	tgggctttct	tttatgccct	18720
taagtaaagg	tttatagttt	tctttatata	agtcttgcac	atttctctgt	agattttattt	18780
ttgcttgtaa	ggtttctctg	gttggtattg	tgtataaaca	ttttctccca	attacatacc	18840
ataattgatg	gttaggagta	taaataaaaag	gtagaatttt	aaaattgctg	tatgaatgac	18900
tgctttgcta	taattcaaca	tcttttcatt	tttatgccaa	tctacttgac	atgttttaggt	18960
taaatgatga	tatctgtaca	gagtaatcct	ctgatccagt	atttgcacat	cttactttct	19020
aacgtccata	gcatagatac	acatcttata	ctgttgagta	catatatatt	taaggattatt	19080
accatagctc	tataatatgc	agcgtgcttt	ggttcaagac	agttgccctg	tgttctctca	19140
cattaacatt	tttttcatca	caaatacaca	ttgaccttta	tcaaattttt	aaaactatct	19200
tgagagaaat	gaccattttt	cttaatctgt	taatgtaaaa	tttttataaa	aatagttata	19260
aataatatta	gcctacatat	ttcttctggt	ctctttttca	actcttagaa	tcagagtatg	19320
gtagtctcag	actaaaccag	gagcttccta	tctgtttctc	tgttcttaag	tcacttatat	19380
aatgtaagga	tgctgtgtat	atctgccagc	taggccttat	atacaaaaagg	cacctatcac	19440
aaccttctaa	aacagcttta	ccacttagag	acctatgtta	aacatatggg	cctttgaggt	19500
aattgccaca	ttcaagctat	aatatgttta	tctaagggaa	tatcttcaat	tctagcagat	19560
gcctaaaaat	atctaaagg	aaacactggg	aattgctgtg	tttggtgatg	ctgctcttcc	19620
tctctctcct	cctcctcttc	ctcctcttct	tctcctcctc	cttcttctct	ctcctgcttc	19680
tctttttctt	catcctcctt	tcttttctta	tttttgaggc	atgatttcac	catgtagccc	19740
taggtaaccc	gtaacttact	atgtatgtag	accaggctag	cctctgtctc	ctgagtgtct	19800
atattaaagg	tgtgtatcac	catatccagc	aacacttgc	ttgagatggg	tagagaaaaa	19860
aaaaatatac	gtaaataaag	atggatgcc	attactaaat	tggtacttcc	agtcacaaat	19920
tgtacctagt	ctaaggccaa	aatagggtat	tttttcttac	tttgcaagtt	ggctccatta	19980
agaggctttt	cttctcttgg	tctcactaga	taggaaggag	agagaggagg	gaaggagaga	20040
aagcggttga	ggagtgggag	gtagtgtgac	cgagaatacc	cagtaggctc	atatatttaa	20100
atatttggtc	cctagttgat	agaactgttt	agaaagatta	ggaagcatgt	cttaggggct	20160
ttgaggtttc	aaaattttaat	gctagacc	gtctttcaag	ggagggggcg	gtctgtctct	20220
ctctgcctgc	tgcattgcaga	gctctcagct	actactctag	tgtcaagcct	gtgtgtcttc	20280
tgctctaatg	ataataaatt	aactgtgaag	aagcctccaa	ttaaatgctt	tcttttatag	20340
ttaccgtgat	catggtgtct	cttcacagaa	atagtaacct	gtggtgattt	taatatgcct	20400
ggaccagggg	gtggcacttt	taggaggaat	ggccttggtt	agaggaagtg	tgtctctgtg	20460
ggggtgggca	atgagaccct	cgtcctaacc	atgtgagaac	cactcttctc	ctattggcct	20520
tcagatgaag	atgtagaact	ctcagatcca	cctgcaccat	gtctgcctgg	aagctgcctt	20580
tgttcccacc	ttgctgcccc	aattaaatgt	tgtacttata	agaattgttt	ttggggggct	20640
ggagagatgg	ctcagcagtt	aagagtactg	actgctcttc	cagaggtcct	gagttcaatt	20700
cccagcaacc	acatgggtgg	tcacaacat	ctgtaatggg	atctgatgcc	accttctggt	20760
gtgtcagaag	acaggacagt	atacccacat	acattaaata	aataaataaa	tataataata	20820
aattcttttt	aaaaaagaat	tgctttggtc	atgggtgtct	ttcacagcag	taaaacccta	20880
acataaccct	gactaagaca	acaagtggag	aaaggtgttg	tgtgacactc	tggatctctg	20940
gaagctcacc	tcagcatgaa	gcttgtcgaa	gcgnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21060
nnnnnnnnnn	nnnatctaag	tacactgtac	tgtcttcaga	cacaccagaa	gaggggtgtca	21120
gatctcatga	cagagggtgt	gaactcagac	ctttggaaga	gcaatcagtg	ctcttaactg	21180
ctgagcatct	ctccagccca	aaataattct	tactagtaac	atggaacaat	caagttttat	21240
tatatgatac	atattaatca	acttataagt	acatgattat	gcacatttat	catatcgtgc	21300
aaccatcact	gctgtcggtt	tgttttgttt	tgttcttttg	aggcccgtgt	tctgtgttgt	21360
tctggaactc	actctgtaga	ccaggctggg	cttgaactca	atgatctgcc	tgccctctgcc	21420
tcccaagtgc	tgaatacaaaa	tgtgtgcacc	accacctctg	gctatcactg	ctgtcttttt	21480

ttttttttta	acagttatatt	atttcgtgca	tgcattgtgtg	tataagcatg	taacgtatgc	21540
catgggtatgc	atgtggagggt	cagaggacaa	ctttcaggag	ttagttcttt	cctcccactg	21600
tgggtttctag	gaaccaagct	caggttggtta	gacttgcatg	gcaagtgcct	ttaccacaga	21660
gccatcctgc	tggccctact	ataggtcctt	atataaaaag	atcatatgcc	gggcaaaaac	21720
caaacaaaaa	ataaacctca	aaaaacaaaa	ggaccatata	atattgtggg	ggagtggatg	21780
aagtcctgaa	cgaatgtgtt	ctgttgacat	gtctgtactt	cagacccatg	ggaattggca	21840
aagccttcct	ctggctcctgt	gaggatgctg	atagtctgtc	taaaaactag	agatcacagc	21900
tttctcctct	ggatgactgt	aaccccagat	tgttcctctt	cagagactgt	ccaccaagct	21960
accctgccta	cttaagctgt	acacaatgaa	tgagctgagt	ttccagggtta	cagcacagta	22020
gacactgtcc	atcagtgaga	gcacagccta	gcctaaccgt	acacatgtct	gctttcttca	22080
cgtttccaga	accaagcctt	gctggataga	gcatatttgt	ctgtttggct	tatttcactt	22140
gataaaaagt	tttcaaggag	ggccagggtg	gggtggcacac	gccttttagtc	ccagcactcg	22200
ggaggcagag	gcaggcaaat	ttctgagttc	gatgccagcc	tggtctacaa	agtgaattcc	22260
aggacagcca	gggtatata	gagaaacct	gtctcaaaaa	acaaaaaaaa	acaaaaacaa	22320
aacaaacaaa	caaacaaaca	aaaagccaaa	aatccaaccc	cccccaaaaa	aaaaacccaa	22380
ccaaaaacca	aaaaacaaca	acaacaaaaa	gtttttgagg	tttaatttat	tgcattgtcac	22440
agaattttcac	tgttttaaaaa	aatggctgaa	taatatttca	ctatccattc	acgtatttgt	22500
aggatttcat	gttgtgtagt	gtttaaataa	aaatagcccc	cataggcttc	tacagttgaa	22560
tgcttagtca	ttgagttagca	gtactagaga	gggaattgaa	gggtgtggcct	tattggagta	22620
ggagtggcct	tgttgcagga	attgtgtcac	tttgagggtcc	cagcaacaag	gttgcctctga	22680
tcacatccaa	agacattcta	ggtctatgtg	atctggctgg	aattcagaca	tgcccttaat	22740
acacaccttt	aatcccaaac	aatgaaggta	aagtttagttt	ataaaaagaa	gcacccatgt	22800
ttgaaagtga	cgttttaatta	agagtgatga	attagagaaa	gatctgctgt	cacagagcag	22860
agaggaaaga	gaggcagcat	aagagggagc	atggcagagg	gagagggagg	aggggttttc	22920
accagggcat	ttgtacagag	acagggttgc	gagctagaac	aggtgaagac	aggaacaagcc	22980
agagaatgag	aaggagccag	gagattagga	cagattgcca	atgttaatag	gctaagcaga	23040
gcatttttagt	cagaaactga	gagaagtcaa	attgaatcag	ttagcttgga	aaggagtttg	23100
agcagcaaca	gctgagttaa	actagccaac	agaatccaga	aagaactaga	aaagatgagc	23160
ttactcagca	gcaaatctca	gaggctaaaa	acatcttaga	cctagattag	actgcatgga	23220
ggctagacgc	ttccagggtc	aggcctaggt	tagcagacgg	agagagtaat	aagccttgga	23280
gacaacagtt	atacacaga	actatgtaca	gacatggata	tgaacctctc	agccactctc	23340
ccagcgtcat	gcctgtctgc	attgttagga	gtcatctagg	aaaggctaag	ggcaggcaag	23400
caacttttcc	agagatggtc	caactgtttt	tgcattggctt	ttgagaggcg	agctctgaga	23460
gggaagggtt	caagagactt	catcccagga	ttgctgctta	attacgacat	gccttttctt	23520
gtcactgtta	tttagtataa	tgactcctga	gcttttagccc	atcctattgg	gcatatttcc	23580
tgcagatcaa	cataaagatg	aactttcaca	aattaatgct	gttttagatga	ataaatgatt	23640
ttataaaaatt	cctgatttga	tttaaataat	tttaggaaga	aagcttttag	agatagttta	23700
gttggttttg	cagaaagatg	taataacgtc	agaatcaaga	atagaatgtg	gctgggcagt	23760
gggtggcagat	gccttttaatc	ctagcacttc	ggaggcagag	ataggcggat	ttctgagttc	23820
gaggacagcc	tggtctacag	agtgaattcc	aggacagcca	gggctacaca	gagaaacctt	23880
gtcttgaaaa	acaaaacaaa	aagaaaagta	agtaaaggct	gcataataaa	gaatacaatg	23940
agctttcaca	actacaccaa	aaagagacat	gcttgggaca	aattttgtgat	caaggaaaaa	24000
tattcattct	agatcagggtc	caaggatgaa	gccacaagtg	tgtgatatga	tgaacaagac	24060
catggataaa	ctgttgtttt	gagcttaaa	ataaaaacac	tgctttgaaa	tttaacttca	24120
acattctact	gtaaactttcc	tttttataaa	ttttatctat	gagataattt	tctaagaagc	24180
ttgtgtctat	aaaggatatg	aaggacagag	agaaaagaaat	aagggtgtggc	atctgggctc	24240
tgctccatcc	acccaaataa	atatgtgtgt	gtgtgtatgt	atgtatgtat	gtttatctat	24300
atgtatgtat	atacatacat	gtgtaggtag	gtatatgtgt	atgtatatata	gtatgcatga	24360
acacttgagg	agttgatgag	acaagtgaga	ggttgggccc	ccnnnnnnnn	nnnnnnnnnn	24420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25320

tattcatgct	tgggtaactg	gccctggctg	attgatgaga	gctgatttcc	ccactgccct	25380
gtggcagggg	ccactgcgcc	cacagggtc	cctcaggatc	ctcagtagag	agctgcacag	25440
ctgggtggaa	gtagagggtc	gcataataaa	cacgatctca	actttatttc	tttaaataaa	25500
aattttattt	aaattttata	cagctctata	taaacgaagg	aactattgaa	ggttcagcaa	25560
ggacctgcca	acggttgta	agggtaatgg	cgatgtagtg	atTTTTTTT	ccccttccat	25620
tttacttcca	tactttctac	attacccac	aactggcaag	tattatttta	aaatgaaagt	25680
aaatagtgac	agatgacttt	gaaggaaaat	tgaatcggtg	aaaagaaagc	tgagagacca	25740
cccgggaagc	ccaggctaaa	tgtaatctgg	gtcaggcctc	ccaggcctgg	ggtctcaaga	25800
tggtcagctg	agggaccctg	gtgaccctct	tgggccagca	gggacgggga	ggagccggaa	25860
gctgagtacc	caaagtgtct	ctctgggtct	caagggtcctg	cacagagact	gtgtgggaat	25920
caaaggatac	aggctagagg	actgaggcct	gacgaaccca	gctatcattc	gtcttagaac	25980
aggaggcaga	gctccaagag	tccaaccaag	aggcaggaag	ctttgggacc	cgagatgggc	26040
gatgggatta	gaaaggcatg	tttgcaaata	ctttcaaatt	tacgatgcac	actcactgga	26100
aacccccacc	ctgggtgtcc	cttccctgcc	tcttgccaca	cccaatagct	gacatcactg	26160
gagaaagtcc	caagaccagg	ctggctggag	ctcctgatag	gttccaccct	cctgcagagg	26220
gccctcgaag	actagcttgc	tgcgccacac	cgccagatgt	ctgtgtcttt	ctctcttttg	26280
cctccacc	tctgtctctt	ctccaacctc	agtggagggtg	cccctgtctt	ctggggaaag	26340
tagaacttgc	cagtgtctac	tgtaatgtcg	tccctgtagg	tgatcatgtc	ccccattact	26400
gggagcaggt	atgcttcaga	tctccctcta	tctcgtgccc	tttcaggctg	tctcagtttc	26460
tctctgacag	ttcctctcct	cctgaatcct	gcttggttggc	atgcgaacag	gctcaatata	26520
ttccatctca	aaaaacaaac	actgggaagg	tggtgagaga	cagagagcat	gggtaatggg	26580
tgccccagct	tggtctggga	gggttaactt	acaatgctct	actgcccagt	agggtagctg	26640
cagttgtcaa	ttaattgtaa	atttcaaaat	agctagtaga	gaggatttta	gatgttccca	26700
atcccaacac	aaagaaatga	taaacattca	aggcgatggg	tatgctaatt	gactgtgact	26760
gatcaccgca	cattgtatac	atgtttttga	aatgtcaggc	tgtaacccat	aaatatgtac	26820
aattaccgtg	cagtgtattca	agataaaaa	tataatttta	aaaagctaaa	aacagaagga	26880
aatagctgcc	cttgaccccc	ccacccccac	aaggctcttc	ctgtttgtcc	agccacttaa	26940
tgtagagct	tctgtggga	gggtggtttt	ggtgtacaca	gacactcctt	cctccctcct	27000
tccccataag	aggagtccac	cctgtcccac	gatgccatgc	agggccacat	gcgtgatatt	27060
aaccagtaag	atgtgagcag	ggatgatacc	tgtctcttat	aacaaacgga	aaaaaaacca	27120
caccaaaacca	aaaacaaaca	aacaaacaaa	caaacacaaa	cagggttggt	ctgtccctgt	27180
gtcttttccc	acataaagtt	aagcacacaa	atgtgccacc	atttatttat	ttgtcccttc	27240
ccccacccct	ccccgagaca	atgtttctct	gtataacagc	cctagctgtc	ttggaactca	27300
ttttgtagac	caggctggcc	tggaactcac	agagacacag	agattcacct	gcctctgcct	27360
cccaaatagca	gggattaaaa	gcattgagcca	cgaactaacc	agtaccccag	agctcttgac	27420
tctagctgca	tacgtatcaa	aagatgacct	agttggccat	caactggaaa	agaggcccat	27480
tggacacgca	aactgtatat	gcctcagtac	aggggaaacgc	caggggccaaa	aaaatgggaa	27540
tgggtgggta	gggaagtggg	ggggagggtg	tggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27660
nnnnnnnnnn	nnnggcagga	tctgtgttct	atgtgcaaca	ctcgatgcaa	gctgtgtagt	27720
gtttggttct	gagcacctga	aggggaccaa	gcaggctgat	gccaggcca	cggttctttt	27780
ctggccccac	tgccccactcc	caccctctgg	catccccatg	atgaacatgg	ccacagatca	27840
cactactctg	ctcctctccc	agatccacgg	agccataggg	tccccagatt	catctctgca	27900
gctaacaagc	tgggcagtgt	cacctccctc	aagggtcctt	tctgtctctg	agcagcagtg	27960
tctcccacag	tgagacactc	atgtccactg	agaagatttg	tagccattaa	attcctgtgc	28020
taaaataact	agggggactt	gtcaatcact	acactcttag	ccccggactt	ctgactcata	28080
gagggtggtg	acagctcagg	gacctgcatt	ctaccaaata	gccatgtgtc	cctgatggag	28140
gaactgcccc	tggacaacct	ctgcagcaac	tgaacctctt	gtggtctcct	agttcttctg	28200
gacaggtgtg	accccagtac	ctagtgccag	gtgagagagt	gctagggcca	cactaagggg	28260
tgacaggaca	aggttgagac	tggtagatgt	ttgggccacc	aaagagaaca	ggtcagtagt	28320
aaaagccata	atggcctgag	ccagcctgcg	agtcctcctt	gcagttggga	cactcttgca	28380
gtgtcctggg	gacctcttga	gggtagcatg	gtcaccaaaa	tcctacaagg	acagatcaga	28440
agtcagttag	gtcaagggaa	cagctctagg	ttctctgtgt	ccctcacgga	cctttttttt	28500
tttttttttt	tttttaagat	ttattttatt	attatatgta	agtacactgt	agctgtcttc	28560
agacagctcc	agaagagggc	atcagatttc	gttacggatg	gttgtgagcc	accatgtggt	28620
tgctgggatt	tgaactcagg	accttcggaa	gagcagtcgg	tgctcttaac	cactgagcca	28680
tctctccagc	cccctctcag	tctgtatgag	acagggcagc	aaaggccttg	tcccagatct	28740
gaggagactc	atgtgaaagt	ccttccctac	ccacccttct	cgaacccctg	aacatcagcc	28800
ccataaactc	tgactcccc	acccccattc	ccttgccttc	actgatccgg	tctcctctt	28860
ccctctggcc	ccacccattc	ttccccagcc	ccacctgatt	gtacctggtt	gtccaacttg	28920
aagagggcag	gcaggggcag	cttctgctgg	gcctgctcac	tcactggctg	tagaaatgag	28980
aaaggagatg	aagaaaaggc	ccttcccatg	ggtcccatc	ttgccaagac	ataggtgagt	29040
ccctttggct	cttcccccta	aacctctcac	ttttgagtac	ctgctggccc	gggagatcca	29100
cggcgctcac	cggagagaac	tggtgagaaa	agggagaaca	gagaactcag	cgttccctcc	29160

tctccaccct	tctggcctct	cccagatttg	ccccgcctcc	ccagcatctc	cttcagcctg	29220
actgaccact	tcccactcag	acctcagctc	tgcctcaccg	tgaacacagg	accttgacgg	29280
caggacaagc	tgagtacgag	gagcccccg	agcagtgcc	tgttcctgta	tccagaacag	29340
ggagtgttag	ttcctacctc	acgctcgaag	gccaagcagt	agactgctat	ccatgggttc	29400
cttgaccgca	ccaggctgcg	gaacctggac	tcaaaacata	gcagctgtgg	acctcactca	29460
ctctgagagg	tggtatttcc	ataagctttt	tttttcacct	gtacatttag	tcttcattct	29520
tttcgtctta	cactgtggat	cagtccctgg	ttcaaattta	aagccctcat	cttgcaagag	29580
gaccttgccg	atctcccttc	atgccttttg	ctttaccctg	tcttggtaat	tcatggcaga	29640
agttcttctc	gtcctccatg	agatgttgag	gacccaaata	agaatctctg	ttaaatactga	29700
gcctgatgcc	tggtcccccac	cctagcaaa	ccacctgacc	tgttggtcat	ttcatccagc	29760
ctttctcagg	tgtccctggg	cctacccaaa	ggctctgaga	gctaactctg	gctggcagg	29820
cagccagaaa	cttctttgtt	gaccaatgaa	tgactggccc	agacaccttt	ggacttacgg	29880
gaactacaag	cctcatccca	cttctgctcc	aagttctgat	ccagggtgct	tcggggaagc	29940
ccagctggcg	gaagggggga	ggctctcagc	ctagagagcc	ttcctttcca	tcctcagccc	30000
cctaccagg	ccttatttca	ggcaccagct	cttctaaaag	gtccttctgt	tatccctaga	30060
cctccacaac	tgtgttcaag	aaccttcagc	cagggcctca	tctccaatct	ggatatatga	30120
ttttctctgc	caagagtagg	cctccagggt	ttggagttct	agagggttct	cctggagctg	30180
cctggacctc	tgtcctcac	cacccaggga	cgctgtgaag	ctgcaggctc	cctgaataaa	30240
ttcatccaga	ccccttgcca	aggtgccagc	tgtctacttc	ctctgctgcc	caagcagcag	30300
gctgcaccac	ccctccatcc	tacctcttca	ggcttcttag	cgcagcacac	gcagcacacg	30360
gtgttctcct	ggaccagctt	gtcctccacg	ctccccaggt	gcagccagca	gggcctagct	30420
ctctcctccc	acaggacctt	tgtctcagc	caacccccgt	tcagcttgtg	ttcagtgctg	30480
gtaaatattg	acctgtacat	ccggttaaac	attgatattg	gggccagaag	accctttccc	30540
atcaaggcta	ccagaacccc	tgctgagccc	tggagaagg	gtttacagga	ccagataagt	30600
gaggaggttg	ggcctggcaa	gccttcta	gatccctcaa	cataggggat	tatccacagt	30660
cagtgaggct	cagagaggct	gtgtggcctg	tgtaagggcg	cagagtgggc	tccagagtca	30720
cagccaaagt	cccaccacca	ccaccaccac	caccaccacc	accaccacca	ctaccaccac	30780
caccaccacc	accaccacca	ccaccaccac	caccaccact	accaccacca	ccaccaccac	30840
caccaccacc	accaccacca	ccaccaccac	cacctcatct	accatacta	anttgaggct	30900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	30960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31020
gaacatctct	ccttttttcc	tttttgagac	aatcttacta	tatttaggct	ccccttgaac	31080
gtgtgatcct	tctgcctctg	cctcccaagt	gctgagggtta	caggcatgca	cagtcacatc	31140
tgcctacaca	atgtcttagc	agcccttagc	agcaccagg	gtcaggaagc	cctcaactgt	31200
cccttttagct	ggcttctctt	gtgaagggt	atgtcttctt	ccccttccca	gcagggagac	31260
ggctcttagc	cccagagcct	tccttctctt	agggttaaaca	gcaccagttt	ttgggtgggac	31320
ctccatttgc	cctctatctc	cctaagcaac	gaccttttct	gctctgactc	tcactctggca	31380
cttggaacc	aagacaaaac	tgcagcctgg	gctgtgtgtc	ctcgcacatc	attcctgtgc	31440
ccccctggag	tcaggcttag	gggaggaaga	cagggttcac	gactcagaaa	agaccactgg	31500
ctgtcctagt	gtgccctcac	ccatcctata	gcacgcacat	gctgatgtgc	cccctccgct	31560
ccatcaccat	cctctcatgt	acacgtgccc	tcctctgcca	gacacatgca	tcactaactt	31620
ttctgacttc	ccagaaaaat	atctgatctg	agaagttagg	agtctgccat	catcagctat	31680
ggctccttaa	attaagttag	acaatccatg	ggacatgaag	ggcaacaacg	agaagactcc	31740
tcgttcttgg	tctactctgc	ttttggcagc	accaccagca	ggaaaccaacc	tggtctctcc	31800
taatccctca	tctatagcag	gtctcccgtt	gggaatttta	gggacctctg	tgttctcatc	31860
caggggcact	gccactcagc	tgtctcagga	gagacccttt	agaacaacaa	agaaatcaat	31920
gcagatttag	gcttcttgtt	tccttctcca	gcccctccca	tcacaggcaa	cagcctccct	31980
tggtctgagc	tcaggaggtc	gatttatcag	agagggtgct	agagaggcac	ctctgggtccc	32040
tctgggtagg	tagcaactga	gacaggagga	catgggtcacc	ctgggcatcc	tctaccagga	32100
agtaaatgag	atacccttgc	agatgggacc	cctgaagttc	ctcccggggg	cggggggtggt	32160
ggtggtggga	gtctaagtca	cagatctttg	ttaccacgtg	gttagactga	ggactgaatc	32220
tgaggtggga	aatctgatgt	gcagggggaa	acacagaggt	ccaatgctgg	ccaagagcta	32280
caagcaggga	cagggtgctag	ggggatgtct	gaatgttcca	ccccaaagca	caggaataac	32340
ggaaatggag	actctaaagg	gcagaaagtg	aggggtgtgca	gcaggggctg	cacaggacac	32400
atgcaaggcc	ctggctgcaa	taactgggtt	ggggagggca	tcattggcta	gccaggggca	32460
ccaggacagt	gatgccatcc	tgtccaaaag	gcagtgctca	agccagattt	ctaggctcca	32520
ggggaggagg	ggtccgggga	gaggggtcaa	gattctcccc	ctctgagtca	aggttggcct	32580
tcccattgtc	cccaaatcag	gaggcacaga	aactgggatg	ttgtggtctc	acatccaaagc	32640
tgagaagaca	agtgggagcc	agtacatgtg	tttcagatta	aacctcagtcg	gagacaaaca	32700
tgttgctcct	cctcctccca	gagccaagct	gccttcaagc	cacatggcag	tgaatatgag	32760
gacagtgcag	gggaggacac	ctctctctcc	actggctcaa	ggacagtttc	aaggggttca	32820
ggctggctgg	ctcatggcta	cgccgctcac	ccccgggaca	gtttgggggt	tttccctcct	32880
gaaatcttgg	aatctgaatc	agcctgagat	acccataat	tgtacctccc	aacaccccca	32940
gaaaggctcag	cctgcagaa	cagaactctt	tggtccccac	ccatccccct	cagccctgga	33000

ggctgaactg	atgggacagct	aagggtccaga	cagtggctg	ctcttgga	gcctgtctct	33060
ttcctttgac	tcagaccact	ccctgccgtg	gcttacatca	ggaggtgcaa	gggctgcagg	33120
agggcagcca	gacccacaa	accagctagg	ctaatgggtg	cttattgttc	gcaagaggcc	33180
atgacctcat	ttgtctccca	gctcttttgg	taagagagaa	tgagaggaag	ctggacagag	33240
aacctagcag	gcctcaggca	gccactgct	ccttgctgta	agggaaaccag	caccgatggt	33300
tctgaaaagc	agcgatccga	atggagtcag	gctgagctgc	aggaagctca	ccttccttgc	33360
tcactgctgg	tggaagcaac	ttcaggaaga	gcccagccta	tgggactata	gctcctccgg	33420
ggtactgctg	agtccagccc	cagagcttag	ctccctgctt	cccaccaccc	accaccacat	33480
cctttcccaa	caccattcaa	aacccagtc	cagcctctcc	tactggtcta	cagtgcagg	33540
ctaataagagt	cctgggcctc	tgtcccccca	attctctctc	ccctctcctc	tgttcacctt	33600
ggttcctaaa	ctgcaggggc	tactataaac	ctacctccac	ttccttgca	ccctcttttc	33660
tgctctctgg	ggtgccctg	ccactccag	tcctcttagc	caggagacct	cttccatata	33720
tgtcttcccc	aggctagacc	aggcgtgcc	ttacctgtgg	ttgcggcagc	ttctctcaca	33780
gcctgcactc	tgaggggctc	caggaagcag	tgaggggagt	agctgcctct	caaccagcgt	33840
ccagcaggct	tcagattaca	gctactcttt	tcttaaagt	acctgactcc	atgtggaatc	33900
tgtgattgca	tcattgtctg	gtgttaactt	taacccactg	ctgcccttcc	gccatgtggc	33960
tccaagacca	cacgttggcc	accctcctct	cccaccacat	ctcccttgga	tcttatctc	34020
tcttcattgg	gaccttcatt	gggacatgat	ggctaacttc	aggggcactt	gggcccagcct	34080
ggggtaggtc	atgagtctga	acttgaacat	ctgaaaggat	tggctgagag	gcaggctgca	34140
tggagagact	gtgagccagc	cggtatggag	atgctgggtt	cttccaggcg	cttggctctg	34200
gctcactgca	ggtgggagca	aggtgattct	tctcccctcc	tcacctggaa	aatgaaggaa	34260
tgggactgta	cctgacagct	ctgaaggttc	caaaggacag	tgggggtggg	actagagagt	34320
tggcccagtg	cttatgagca	ctggctgctc	tcgcagagga	cctgagttct	gttcccagct	34380
cacatagcaa	ggactcgaaa	ctgcttgga	ctcagctcc	agagaatctg	acgctatctg	34440
ctnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34560
catgccttta	atcctagaac	tcaggaggca	aagcagatgg	acatctgaat	ttgaggcccg	34620
cattgtctac	atagcaagtt	ctaggctagc	caggctacatc	ataagaacct	ttctcaaaaa	34680
ataaataggg	ccagttggca	aaatttagct	tgccctccta	acacaagaac	ccaaagtcaa	34740
ccccagcaac	catgtaaaaa	gaagcaaggt	gtggtggcac	ttgcttgtag	tccagcattg	34800
tcaaggtgga	gacagacgga	tccatggggc	tcactggcca	gccagctagc	tggtctactt	34860
agtatgctcc	cagccagtga	gagactgaaa	aataaataaa	taaataaggg	gttaggaaga	34920
ggtaacatgg	tggctcagtg	agaaaagata	cttgtcatac	aagcctagca	accctcaatt	34980
ttcagtggcc	actaaagggtg	gaaggagaga	accaactcca	aagaattgtc	tcctgacagt	35040
tttatgctgt	ggtacacaca	cacgcacaca	cacacttggt	acatgcatat	gcatacaatt	35100
aataatttaa	aatgttatgt	gtatgggtgt	tttgcctaca	tgcataatct	tctgtgcacc	35160
acatgtgtgc	aatgcctgtg	aaggctagaa	gaggacatca	gatccccctg	gagttacaca	35220
gggttgtag	ctaccatgtg	gattctggga	acaaaacct	gggttttcca	aaagaggctc	35280
ttaatcactg	agccatctct	ccatccccct	aatagtatat	atttctgggg	ctggagagat	35340
ggctcagagg	ttaagagtat	tgactgctct	tccagagggt	ctgagttcaa	ttcccagcaa	35400
ccacatgggtg	gctcacaacc	atctgtaatg	gaatctgatg	ccctcttctg	ctgtgtctga	35460
agacagctgc	agtgtactca	tatacataaa	gtaaataaat	aaaccttttt	tttttttgtt	35520
ttgttttttt	tgttttttga	gacagggaga	cagggtttct	ctgtatagcc	ctggctgtcc	35580
tgaactcgc	tctgtagacc	aggctggcct	tgaactcaga	aatccgcctg	cctctgcctc	35640
ccaaagtctg	ggattcagg	tttgcgccac	cgccaccacc	tccaaggctg	ctgctgcggc	35700
caccaccacc	ccaccaccac	actacctgac	tatttaactt	ttaaaggcag	ccatctcatg	35760
gaaaatgaca	cctagcattg	tcctctgggtc	cctacatgac	cccatgtgca	aacacatacc	35820
tgcataaaca	cacataaata	cataagtaaa	cttagtctgg	ttgttttgga	aatgtgctat	35880
ggtttggatt	gtgtcccccc	aagggaaaaa	ttgggtccca	gagtagtact	attggaagag	35940
agtagaactg	ttaggggtta	ggcctgggtg	gaagtggcca	tggctagaga	gacatgccaa	36000
ccaaggggaa	tctctggctt	catcttttcc	ctttgctttc	aggtcctaag	acagtcacaa	36060
ggctgcttca	ccacatgccc	agaagcaagg	ggagcagtc	tggctgggac	cgctaagtga	36120
gctgttgatt	tccccagata	ttttagtag	taagagacag	gtgaggaacc	ccacagcaag	36180
tgttagtaat	tgtgtgtgga	ggtgccctcc	ggggacgggg	gccctcctgg	ggcaggacgt	36240
tcctcttctc	catccacctg	cactccgaga	acaggaaatg	gtgacttttg	cagagcttaa	36300
gcagagcccc	ttcatgttac	aagtatgtaa	attcataagg	accagtttct	ctccatatga	36360
aacagcttca	aacaggagaa	ggaagaagca	aacattaagg	aaaagctctt	ttattgcaga	36420
ggctacagtg	aagctaccgg	ccgccttctc	ggaattgata	atcagcttcc	ctctgggggt	36480
tctgtagagc	actgagacat	taagtactac	tggggtccag	gattctgcct	atgaagagga	36540
gggcccccg	gtccgtgtcc	ctcagaacaa	agaggaaagg	ttgggttaagg	tgatagtcta	36600
gcgggaaggt	gaggcgagcg	ggctggaggc	ctgggctggg	gctgcttctc	gccccctctt	36660
cattccactc	gaaagcagcc	ctgtgttcca	cttgggtgag	cttcacgggt	ttgccagtaa	36720
tcttgctgaa	gtcgggtgat	tcaaacaacg	actgtagctc	tgtggagatt	cagagattcc	36780
attaacacca	cacacacaca	cacacacaca	cacacacaca	cactccctgt	ttgtgtaggc	36840

tgatttttcaa	gaaagcaagc	tagaagtggg	gtacctcaca	gtgacttgtg	agctatgagg	36900
cactctgtga	caggctcagt	gacctacctg	agaacttata	gccaagatgg	ctgaagccag	36960
acctggcctg	agagaatgtt	ttgggctgtt	ataggacaca	tagagataca	cacacacaca	37020
acacacacac	acaccaagga	ctgagtctaa	tgggaggtgg	ttcttcattc	ccctcccctg	37080
taatgggtgc	acatgttccc	tgagccaccc	tacaaagaaa	gccacaggac	tcagttctgt	37140
cagcaaggtg	gcaggctcca	agactcagcc	ccgagcgcaa	agtggccttg	caaacatact	37200
catgtcctgc	agagacttgg	taagttcgcc	ttcgaagctc	agcttcagct	tggggacagt	37260
cagcacagct	tggatagtct	tcagttctcg	gtcgaatgtc	tgaatgaact	cagaggtgag	37320
gctctcttct	atcatgggtc	agttctgggt	cacggtcagg	ggcaggaaga	agatgatgct	37380
ctacttctct	gtcaagggca	gctgggcaat	ctaaccacaac	agagatgcgc	acaggttagt	37440
tgtgagccag	aaaaaacaaa	acaaacaaac	aaaaaaaacac	caacagctgc	cttcccctct	37500
gctgtaacgg	ggccccagcc	ttgtgctccc	cagcctcagc	ctgggctgta	ggctactggg	37560
tactggcagt	ccttccatga	gtagggagtt	ttcttctcag	cctaaaaccc	acagaagttt	37620
aatgaacaca	cgtttgtttg	tggttccgct	acggtttcta	ttgtgataaa	acatgactga	37680
aagcaacttg	gagaggaag	ggtttatttc	atctgacaat	tcgcagggtg	tcttctcatc	37740
tctaagggcg	ctcagggcag	gaactgaagc	ggaagccgtg	gaggaacgct	gctttctggc	37800
ttgtctcccg	tggcttctta	gcctgctttt	ttatgtctatc	cagaaccact	tgcccaggag	37860
tgacactgcc	cattgtgggc	tgggcccccc	cacatcaatc	actaatctag	aaaatgacc	37920
acgggtttgc	ccagaggcca	gtctgggtgg	ggcattttat	caattgagtt	tcacccttcc	37980
aaatgactct	aacttgtgtc	aagttgacca	cacgaatcag	ggcctggttc	ttaggagctg	38040
aagtggatg	tccccagag	actgcctgcc	agcactgctg	accatttgct	ttgtatagag	38100
cattgaacca	gaaatgaaca	ataaaatgga	tcctttgaac	agatgtgttg	atcctagggc	38160
ctgtggacac	agcgactggg	cttcccagag	ccccatgga	atcannnnnn	nnnnnnnnnn	38220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38340
gagagagaga	gagagagaga	ggataaaagg	ttagcccagt	ggtggtggca	cataccttta	38400
attccaacac	ttgagaggca	gaggcagggg	gagctctgtg	ggccagtttg	gtttacagag	38460
taagtttcag	aatagccagg	gctacacaga	gaaaccctgt	cttgaagaga	aacacacaca	38520
cacacacaca	cacacacaca	cacacacaaa	taagatcttt	aagaagaaaa	gaaaggatag	38580
tggggaaaca	tctgagcaga	ggaagaaatg	gggtgcgcag	gacacccacc	ctcagaggag	38640
gccctcactg	gaggtgtctg	cacaggagaa	cacttgcaat	cagcttgccc	tagggcgta	38700
gaactcagaa	ttcagtttca	aagcactgac	aggagcagtg	actggggacc	ccaggttgaa	38760
tccccctttt	atctaaaatg	agtaagaacc	aaaaaaaaca	aagtgttttg	gatttggaat	38820
ctgggttatt	tgcactctaca	aaagaggtct	tggggaggaa	accagtttct	acccttgga	38880
ttcatcagtt	tcctataccg	ctgactacac	aggggctgaa	ggtaatctca	atgttttcat	38940
aactggtgtg	gtgctacttg	ctcatgatcc	caacacttgg	aaggtaggtc	agaagttcaa	39000
gagcagtctt	gactactcag	tgaatttgag	gctagcctgg	gctacatgaa	aactcataaa	39060
acaataaaaag	aaaagaaaag	gtggcgagtg	aggtggccca	tcaggtaaaa	gtgccaaaca	39120
cctcgctga	aagcctccac	acggaagggg	aaagccagct	tctacatgtg	gtcctctgcc	39180
ctccgcatgc	accatggctc	gtgcaccccc	acacccacc	acccacccac	ccacatgaca	39240
taaatacttg	taatgattag	tttctgaaga	acaatatatt	cgttgatctt	gtttagggaa	39300
caaagttcgt	gcacattgac	ctgtcgaccg	tgtagtacgg	gatccgctcc	aggaagctaa	39360
agatttttgg	atgcttcatg	ttgcagctac	tctgatgaac	agtgttgctt	tctgggcccag	39420
gtaatggtgg	catatacctt	tgatcccagc	acttgggagg	cagaagcatg	tagatctctg	39480
tgagttcgag	atcagcctgg	tctacagagt	gagttccagg	atatccaagg	ctatacagaa	39540
aaacccctgt	ctctaaaaat	cactaattta	aaaaaaaatt	ccttttctaaa	cctatataaac	39600
aaatgttttg	taggctgcct	taacaaaagc	caatggccat	tcagagaagg	ctcaaaagag	39660
aaagtttagg	ggactctaca	agcatcctca	ggaaggccac	agaaagcaga	gcctgggcca	39720
gtgagacttt	gcagtgggca	aggttcagct	ctttctgtag	gaagaagaga	gtcaacagtc	39780
agagtcacag	tttccataaa	acctgtgcag	ggcctctagg	caaagccctg	tgttaggggc	39840
aaaggcattt	gcagtctaag	cccggtgaca	tgagctcaat	ccttggaacc	caggtggaag	39900
gagtgtgctg	actccacaaa	tttgtcctct	gatctataca	tgtatgcacg	tgacgcacac	39960
ctcacatata	tgtccacatg	cacacatgca	tatacatgcg	catgcgtgca	cccacacaca	40020
gggtcaaaaag	cagcaagaga	tgccctgtga	aaaacgtctc	attcagttctc	ccatcatcca	40080
gtgccacact	ctgagcacag	gtggtactga	tatcgtttct	gattgatcga	tcagttgatt	40140
tgagaccccc	cctcactatg	tagcccaggc	tggcctggaa	ctcacagtga	tcctcttgct	40200
tctgcaagat	gagcccatca	tgcccaccat	gttattgaag	caataccatg	ctctataaag	40260
caaacctagg	caggcaggat	ggtggactcc	tgtaattctca	ggacttgaaa	agtagaagg	40320
agatgaggag	ttcacatcaa	tctcccgtat	gcgttggagg	ctggagtggc	tgttccctgg	40380
gcgcttctgc	cagcacctga	ccaatgcaga	tgcagatgct	cacagccaac	catcagactg	40440
agctcgggac	cccagtggag	gtgctggggg	gaggactgga	ggagctgaga	cgggattgca	40500
agcccatagg	aagaacaatg	tcagctggcc	aaaccaccca	gagctcccag	ggactagacc	40560
acgaaccgag	gactgcacat	gaagggatcc	atggctccag	atgcatatgc	agcagaggac	40620
agccttgctc	gacagcatgg	gaggggaggg	cattgggtcct	gtggaggttt	gatgccccag	40680

tggttgagga	tgctggagcg	gtggggcagg	agtgggtggg	taggtgggga	gcaccttcat	40740
agaggcaaaa	gggatggggg	agaaggcaga	tgggatgggg	gggttggtga	ggggtaagaa	40800
agaaaaaaga	tgtctctgaa	agtataaagt	acttgtcact	aagcatgagg	atatgagtca	40860
acccccaaagc	cccacagggt	ggaaggagag	aaatgagtcc	cacaagttat	tttctgatct	40920
atacgtgcaa	tccatggcat	acgcagaaac	gcaaagacag	acaatgagtt	gggtgtggtg	40980
gtgcacatgt	aattccatca	ttcaggagac	agaagcagca	gagttgttgg	aaatctaagg	41040
ccaacctaaa	gacctacacc	caaagaagga	caaactataa	ggaaaaaggt	ggtcgaccaa	41100
tgtaacatta	aagttagaaa	tctctcttca	cactgtgtag	atactgtaca	aggaagagaa	41160
aaggcagcca	catcaaaaaca	gtgtaaatac	acgagaaaaac	cagaaacaac	tcaagagaag	41220
gctgcagggg	cctgaattct	gttctcagaa	cctgcacaa	gccaagagaa	tcaaaactgt	41280
ctgtaactcc	agctccctgg	gatccaacac	tctttcttgg	cctccatcag	catcactcac	41340
agggtgtgac	acatacacat	caataaaaaat	caaaaccagg	gatgaagggg	tagggaggtg	41400
catgtggatc	tgggaggagc	tgagaggcac	tgggtgaata	caataaaaaa	tttgggtgcat	41460
ggtggtgac	gcctttaatc	ccagcacttg	ggaggcagag	gcaggcgaat	ttctgagttc	41520
aagaccagcc	tgttctacag	agtttagttcc	aggacagcca	ggtctacaca	gagaaacctt	41580
gtctcaaaaa	aacaaaacag	ccgggcggtg	gtggcacacg	cctttaatcc	cagcacttgg	41640
gaggcagagg	caggtggatt	tctgagttcg	aggccagcct	ggtctacaaa	gtgagttcca	41700
ggacacagag	ggctacacag	agaaaccttg	tcttgaaata	aataagcatt	tgttgctggt	41760
acagaaaact	ccagcccagt	ttccagcaca	cacagggtga	ctcacaacat	cataactcca	41820
cttccagggg	atccaatgcc	ttcttctgac	ctctgtgggc	accaggattg	catacagtg	41880
acagacatgc	acataggcaa	aacactcaca	aaataaaata	aatctagcaa	aaaaaatttt	41940
aactaataat	ttaaagaaaa	aaataaggaa	gccgggggtg	gtgtcgcacg	cctttaatcc	42000
tagcacttgg	gagacagagg	caggcgggatt	tctgagttcg	aggccagcct	ggtctacaaa	42060
agtgagttcc	aggacagcca	gggctacaca	gagaaacctt	gtcttgaaat	aaataaataa	42120
ataaaaaata	aggccaagta	attcttggaa	gaatcccaag	gggacactaa	gtgtatataa	42180
aggcgttcca	tagggctagg	aatgaggctt	agcgagagca	acttcgctgg	tgtatgaaag	42240
tccctcagct	gcatgtggta	cctttaatct	aggctctccc	gaagcagagg	cagaaggatt	42300
tctgtgagtt	caaggccagc	ctgggtgtaca	tagctagttc	caggacagaa	agggcgatat	42360
aatagaaaca	tctacctag	agcccngcca	aanaaagggg	agacctgaga	ccagagagat	42420
gactcagtgg	ctaagagcat	tgactgctct	tccagaagtc	ntgagttcaa	ttcccagcta	42480
aaaattttat	taaatgttta	ttacttgtat	tattatttaa	atttaataat	ataagtaaat	42540
gggagcctag	gtttgagttc	ccaaatcacc	aagaaaaaat	gttatcattg	ctaataatca	42600
aattaagagc	ataagaactt	cttttttaag	aattcttatt	tattttatgt	atgtaagaac	42660
actgtagctg	tcttcagaca	caccagaaga	gggcatttga	tcccattaca	gatgggtgtg	42720
agccaccatg	tagttgctgg	gaattgacct	caggacctct	agaagagcag	tctgtgctct	42780
taagtactga	gccatctcta	cagctcttat	cagggttgata	aaatttaatc	tcgtggagcg	42840
ctgagaccaa	gaactaaagc	tgggagattg	aaaaatgcag	accaccaagg	ccctgctcat	42900
ttctccagtt	ctgatcagct	cccgtaccag	gggtctaaac	aggcctgtgt	ctgcttccct	42960
gagtagacca	gaggcccat	ctaaacagcc	tgccctgcagc	agctcctctc	tctaggtgga	43020
cagatgggaa	tttcagacca	atgtcatttc	ccaggacatc	aacacagcag	ccaaatttat	43080
tgggtgctgtg	gctgccacag	ttgggtgtggc	aggatcaggg	gctggcattg	gcacagtgtc	43140
tgattattgg	ctatgccagg	aaccagtctc	tcaagcagca	gctcttctcc	tatgccatgc	43200
tgggggtttgc	cctgtctgag	gccatgggac	tcttctgttt	gatggtcgcc	ttcctcatcc	43260
tcttcgccat	gtgaggctcc	ctgggggtcac	ccagccgtcc	ctgctgcctt	gactccaatgc	43320
cagttccgtgt	gtggagttct	actgagattt	accattaaac	agcaacgttt	ctctaaataa	43380
ctattaatta	attaattaat	cacgtgacaa	ccccagcgtc	catatgggtg	tggaaaaatga	43440
ggaactctac	ccatcataca	tggcgactat	gaagaacaat	gtgacagaaa	atgctaaca	43500
catgtgtgac	cgcattgcac	agccctgact	gctaaaagtg	gacaagcccg	aagcgaaagc	43560
ccaatgttct	acttctaaat	gcatgcacca	aacgcctccc	acaggaccag	aggtgcagct	43620
ctgatagggt	ccttgccctgg	catgcatgaa	gctgtggaca	cgaggcatta	ttcgcaagaa	43680
cattctagct	gtctggaggg	ccctcaatcc	actgtgttgc	cgctgttcca	gcacctgtgc	43740
ctcctggggc	tgcacctgaa	aaaggggact	gcttaagagg	gctcctacca	agcctaactgc	43800
cacagatgca	tgatgggaaa	gccttctgga	agcaactggc	tgccaaaggc	tctggacaag	43860
agatcaccct	ctactggaaa	ggtgggtttca	gtctaggttc	tgtgggattc	caggaaatta	43920
gacaacactg	gcagtccaac	agacagacga	tctaaacttc	caaggcacag	ctggtagaac	43980
ttgtgcgga	accagacaac	aaggtagcag	ctactcccat	acaacataca	aaaaagcaga	44040
gagagagtca	gagacagaga	cacacagaga	gagacagaga	gagagtctaa	agagagtcag	44100
ggctcagga	ctgagggtat	agtctactgt	agagcatttc	cccagcatat	acaagaccct	44160
ggattcaatc	tgaacacagg	aaaaaagggg	gggggggact	tcgattatct	caaattctct	44220
tttttgtgac	acacccctaa	agtcactgcc	tacttccctc	accgccatga	agtaaagagc	44280
tgtttgcgct	tatgtctaca	cagtctcggc	tcccacttcc	tcctccctcc	tgtctctgtg	44340
ctcatctcct	ctgaaaccac	tgacagcaagt	gacttgtgtt	gactgccaca	cggaactct	44400
cctcagtagc	aggcagcaga	gcagagctct	gtcttctcgg	agcttcttct	ctcttgtcgc	44460
cattttctcc	cacccttaag	taccctatct	tctctgtctc	tgttgttga	tccttggacc	44520

cttttccttt	ctatgaacaa	aatatctcct	taaaggatct	cttctagttc	agggtccccc	44580
cgccccact	gtggagaaaa	cccagggcct	tgcacatgct	cagcaggagc	tccatccagt	44640
ctctagctcc	atgacctaaa	gcactctctgt	gctgtcaaat	atacacttcc	agcccttacc	44700
aaaatattca	gtcaactcct	tgccattcaa	aatggatgac	ctcaaagcca	gagtcagcgg	44760
tgctatgact	cccagatcca	tccacttgggt	agcccaggaa	tgaactcann	nnnnnnnnnn	44820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	44880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnngg	ttactgggat	ttgaactcag	gacctttgga	44940
agagcagcca	gtgttcttaa	ctgctgagcc	atctctccag	ccaccaccac	caccaccacc	45000
accaccacca	ccaccaccac	caccccacca	ccaccccgcc	ccacctgctc	attcctgatt	45060
ggttggttag	tttagtctgt	gagacaggag	ctgtcccttt	tctatagtgg	aaggtgaata	45120
agaaactcct	gaaagtgaag	gcctacaaaa	cgagcaactt	tatttggttg	aaaatactgt	45180
aaatgtgaca	tgtaaatata	tgctaaaata	attcggttaag	tcagtgaaca	accttaaaac	45240
acagtctgta	gcctgaatta	cagacacgac	acgagccatg	acagaggctg	aaataagacg	45300
cctttgcaag	gagaagggca	gaagcttcca	tccttgctag	caaactctttg	ttccaagctt	45360
tatcagattt	tattgctttc	tctttctggt	tttctttatc	atatttggtt	atttggtggg	45420
ggaaagccta	atcttcatag	cccatgtatg	tgagcacttc	agcatatgtg	tgtgaacacc	45480
aacagcacac	gtgtatgagc	accacagcat	agggtgtgtga	gtaccacagc	acatgtgtgt	45540
gagtagcata	gcacgtgtat	agagttcaga	cgagcaactt	gagagtccgt	cttttctctc	45600
tactgtgtag	gtctcagggg	tggaacttgg	gctcagcctt	ggtggcaagc	tcctttatcc	45660
acagagtcat	cctgccagcc	cagctttctc	tttttctctc	tggtatgtct	atccactctg	45720
ttcaaggcta	actcactgac	tctgagttat	cagaactgct	tgtgagagca	ggagtaactt	45780
tggacatctg	tgctggtagg	aacaccatcc	ccactcggct	tggatgacga	aggggaaaaa	45840
aagcatcacc	aaggagttcc	accacctcaa	ccagcaaata	tttacctect	atacatggat	45900
aggtaggggt	ggtgagcctt	gtgatttata	gtaggactct	catgggagtg	attacagctg	45960
gtctactcca	tgaccaaaa	ggtgacggtg	gctgacaaaa	aagaaacagc	tacacctggc	46020
tctagttttc	tttctttctt	ttttcttttt	cttttaccct	acggtactaa	ggattgaacc	46080
caggaatgca	agagctctgc	caagtgaact	acattcccag	atctgttttt	ccatttcttt	46140
ctttcctttt	agattttatt	tcgattttatt	tgtctatggt	tgtgtgtatg	tgtgtatgct	46200
tatgtgtacg	tatcagtatg	ccatgggtat	acagaaacct	gagaaggcca	gaagagagtg	46260
tctgggttac	aggagtttcg	agctgtcctg	tgggttctca	aatgcagcag	caccaccacc	46320
aatcaccccc	acccccaccc	agccttcgag	ttcaattctc	atcatcacaa	aaacacacac	46380
acagacaagg	gcctgcaaga	tggtctcagca	ggtaacgaag	ctcgtgtcat	aagcctgaga	46440
acctgagttc	actgtctgga	accgcgtaaa	agggggaagg	gaagaatcaa	ctctatgatg	46500
ttgtcctctg	ctctcccat	gtgtgccatg	gaatgaacag	ccctcccaaa	cacacatcaa	46560
gaataaataa	aactaaaatt	agcttagtaa	cttttatggt	gaaagtgggt	tttacatgcg	46620
tgggcaacaa	taacaccgag	agtagaaaagg	caagcatgta	tgtcactgaa	cagcattgaa	46680
gaaaaaacia	acacatttcc	tgtacatcgt	tctgggagtc	tgagttaggg	tttctatgct	46740
gggataaaaa	cacctgaca	aaaatttaacc	tggggaggaa	gctgtttatt	tcagctttta	46800
tgtctacaac	atgacctgtc	accaggggaa	gtcagggcag	aaattcaaac	agggtcagaag	46860
cctggaggta	gaagctgata	cagaagccat	ggagctgctt	ctggcttgct	ctagcacaca	46920
ggagtactag	cctaggggct	gtactgccc	cagtgggcta	ggctctccca	cagcaatcat	46980
aaatttagaa	aatgcactac	aggtttgcac	acaggccaat	ctggtagggc	cattttctca	47040
attgaggttc	cttcttccaa	aaggacttta	gcttgcatga	tggtgacata	aaaactagcc	47100
agcatattgg	gattatagat	attctcataa	aaaaaagaca	tttagattcc	caataaacac	47160
catattcaga	aattaaactca	atgtgaacca	gaagctctga	aagtaagagt	taaaactatg	47220
aaaaattctt	acaaccatcc	ataacaaaaa	tctgatgccc	tcttctggag	tgtctgaaga	47280
cagctacagt	gtacacacat	ataaataaat	aaataaatat	ttaaaaaaat	atatgaaaaa	47340
tcaggctggg	gagatggctc	agtgggtaag	agcacccgac	tgctctttcg	aaagtccaga	47400
gttcaaattc	cagcaaccac	atgggtggctc	agaaccatcc	gtaataagat	ctgactccct	47460
cttctggagt	gtctgaagac	agctacagtg	tacttacata	taataaataa	ataaatctta	47520
aaaaaaaata	aaactatgaa	gaactatgaa	ctacaagaag	tcaggaatag	ggctgggggt	47580
gtaacccaac	agaaaaacac	ttgcctggcc	tgctgtttgt	ctctagcacc	accaacgtag	47640
aaagagaaca	gcagaggatg	agggcatcct	gacttgagtc	aagtgacaag	tgataatcct	47700
cgagacacca	aatcaccaat	gataaaagag	atcaacaagt	tgggctttat	ctgaataaag	47760
agctgtgtcg	ttaaataacca	cgcaggaagt	gaagaggagc	tgagtctggg	aacacaggcc	47820
tgaaatccaa	gctactgggt	ggactgaggg	aggacaacag	ctagctcaag	gcccacctgg	47880
acgccagagt	taactcagag	agcagcttgg	gtagctctta	tgagactctg	cccaggccag	47940
tgcacaggag	agatggctca	gtggtttaaga	gacttactcc	tcttgcaaa	gacctgagtt	48000
caattcccag	caccacagtg	ggcacttaca	atcatccata	actttagttt	caggggatcc	48060
aatgcccttt	tcacagtacc	aggcatgtac	acagtgcaat	tacatacata	catgcatgca	48120
tgcatacata	cacaggcaaa	acttacataa	aatactaagc	agataaatct	taaaagaagc	48180
cgggcgtggg	ggcgcatgct	tttaatccca	gcacttgagg	ggcagaggca	ggtgtatttc	48240
tgagttcgag	gtcagcatgg	tctacagagt	gagttccagg	acagccagga	ctacacagag	48300
aaaccctgtc	ttgaagaaaa	taaaaaaaa	aaagaaaaaa	atcttaaaag	aaaaggagag	48360

gactggagag	atggctccac	agttaagaac	acttggttctg	aggtctacag	agtgagttcc	48420
aggacagcca	ggactataca	gagaaaccct	gtttcgaaaa	accaaaccac	aaacaacaac	48480
aacaacaaca	acaaaaccac	ttgttctttac	agaggacttt	ggtttgattc	tcagaatcca	48540
catgatgggt	cacaaccatc	agttgcaggg	atccaaggctc	ctgtcttctg	tgggcaccag	48600
gcatatatgt	ggtgtacata	catgtataca	ctcatataca	taaaaataaaa	agtttttaaaa	48660
aggaggctgg	gtttgttagcg	cagaggtaga	ggtaaaaaaga	ctctagcttg	tttaatgttg	48720
acatgaaaaa	aaaaagacat	ttagattcct	gcatcacacc	atatccaaaa	attaactcaa	48780
tgtgaatcat	aagctctgaa	agtaagaata	agcctagtat	gcactgtaag	gctctggggt	48840
cactccccag	cactgcaaaa	gatcatgaaa	ccagaaatgc	agatcctctg	aaccacagca	48900
tgggaatgta	actcagccga	tgcagtgtct	acctgtcgta	tacagagcac	aggataaatt	48960
gattgtgggtg	gtgcatacct	ataagctcac	tacgtggaaa	gtagaggcag	gacgacaaa	49020
ggttcagtgga	catccttggg	cacatagaga	atlttgaggcc	agtctgggtct	gctggtctat	49080
ttggaatgct	gtctcaataa	ataaaagaaa	gaaagaaaaa	gaaaagaaga	agtcctatga	49140
ttgtcttaac	ctctgacctc	tgtgttcatc	aagtctcctc	ctcaggaact	cactgggtcat	49200
cttgtgaaaa	cctaccccag	agtctctgtt	cagaggaccc	aggctccagc	tgtgggttacc	49260
acataggatt	tttatactag	aaaaataaaa	tgaataagta	tgtatTTTTT	aaaaagggtgc	49320
agagctggat	atgggtgggtg	ctagttatag	catccagaac	tgagacagga	tagccatgag	49380
gttgagaaca	gctagactat	acggctctcaa	caaacaaaag	taagggatct	gagttagatga	49440
ggttttaatt	tttttctttg	tgtttgttac	ctaactgtga	tgggtgtttt	gaatacatgc	49500
atgtctgtgt	atcacttgtg	tgcctgaaac	ccaaggaagc	cagaggaggg	catcgggtcc	49560
cccgggaagta	ttattacaga	aggttgtgag	cagccatgtg	gggtgctgga	atcaaactctg	49620
aaagagccac	ctcgggctgg	agagatggct	cagtgggttaa	gagcactcaa	tggctgctct	49680
tccagagggtt	tggagatcaa	atcccagcaa	ctacatgggtg	gctcacaacc	atatgtaatg	49740
ggatccgatg	ccctcttctg	gtgtgtctga	acacagctaa	agtgtactca	aataaataga	49800
tcaaaaaga	aaaaagaaa	agccacctct	ccactctccc	tttttaaaat	cctcttgctc	49860
ctgtccctta	atgttaataa	cacagggtata	tgatactatg	ccttgtttat	gaatagaaaa	49920
tacacgtgct	aaagcaagtg	tgaaccttaa	atacattatg	ctgagtaaaa	ggagttagtt	49980
gcacacaaga	cttttctgct	caagagtatc	tgtatgaagt	attgaacatg	tgaactctga	50040
aatcgggagc	tgaggaagat	atggggagtt	ctaattggcta	caacatttct	ttttggaatg	50100
atgaggatgt	tctagaactc	aaaaatgggtg	ataactcagc	atatatacta	aaactcattg	50160
aattgtacac	gttaaatgaa	tgcaataaaa	cttgtctcag	taatgtgggt	tagaagatgt	50220
acagacatgt	gtgtgtgtgt	gttaaaacat	ttctttggcat	ggcaataaaa	atacagtttt	50280
agccagggtgg	ttgtggctca	aaaaataatg	ataataacaa	taataaaaaat	aatgaaaaca	50340
gaggctggag	agatggctca	gcggttaaga	acactgactg	ctcttccaga	ggtcctgagt	50400
tcagttccca	gtaaccacat	ggtggttcac	agacatctgt	aatgggatct	gatgccctct	50460
tctgatgtgt	gtctggaaac	agctacagtg	aaagtcattg	caaggacttt	acaatagtga	50520
ccatgataac	attgaagcta	gacttgctac	tactgtctgag	tgtgtctgtc	ggctctttct	50580
aaggagtaat	gttagctttt	tgtcctaaat	ttgtttcctt	cctttcctct	ctccctctgc	50640
tgttttttct	tacccctctt	ttactttgct	ttccctctct	atctcctctc	ttaacagagt	50700
tgtcctatgc	agcccaaatg	ccatcttctt	gcctcagcct	ccccagtggt	gaaaaatact	50760
ctttccacag	gttatgttag	gagactggag	tctgtctcagt	cggggaggga	gcctgggtca	50820
agttctgagc	tcaattcctt	ttctttcttt	ctttctctct	ttctttcttt	ctttctctct	50880
ttctttcttt	ctttctttct	ttctttcttt	ctttctttct	ttctttcttt	taagacaggg	50940
ttttctttgta	taccttggtc	gtcctggaac	tcactttgta	gctggcctgg	aattcagaaa	51000
tctgcctgcc	tctgcctccc	aagtgtctgg	attaaagggtg	tgcaccacca	ctgccagacc	51060
ctgggctcaa	ttcttaacat	tgtggagaga	aaagtattgt	agctgttctg	gccacctgga	51120
attactttgt	ttctgatctt	ttgtgtcagt	caaactcctc	tcatccatct	ttcctcgtca	51180
ggctataata	tagactctcc	ttgcaatact	tggaaatgct	ctacagtcag	ctacatcctc	51240
agtcctgctc	ctatatTTTT	tcctaagctt	ccttctaagg	tctttattgg	tttatgattt	51300
acacagaaca	tttttttttc	ttgtctatag	catgcgttag	agtgatcggt	gccagataga	51360
ggaaagagaa	atgagagaa	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51480
cagctactga	ttcctcctcc	tcctcctcct	ttcctcctc	ctccccagcc	tcatgctctg	51540
ctcatcttgg	acttctgcgc	atgtcctcag	ccagacctt	ctgctcttgc	ttctcctctc	51600
cccagcagcc	ccccagttct	cttctctgaa	cttctgaggt	actctccatc	acctcctttg	51660
gctcctgctc	tgattgggtg	cacctgctgg	ataggcttgc	tcctgactcc	actgttctgt	51720
tctcaattag	ggacctcac	cctctgatat	accacacatt	tccctagtgt	ctccacctcc	51780
cacccccacc	ctatacgcac	atacacactt	agctgcacata	ggatcctaca	ccagggactt	51840
cttaccttcc	taatcctccc	caccggacac	tgcccaggga	cactggggct	ccagagggct	51900
attgccacac	ggacacacag	gagatctcat	caaggagatg	tgccctaccc	agagggtagc	51960
tctcaccatt	cacaagcaca	ccacttctgc	ctccagcttc	tactctctcg	caggaagtga	52020
ccagcccggt	gccaagtatc	cccaactaca	tccccaaaat	tctcagacac	tgccagcctc	52080
cagctgtcag	cctggccccg	gctggcgggc	gcctgtcctc	ggcatagcga	ctagggtgta	52140
attagaaacc	cgctagctcc	ctaattgcca	gttctgagct	gtccttggtta	ccggctgccc	52200

gaggcacaca	tagaggaaaa	ggctgagagc	tgagccaggc	tggcatggag	gtagccctag	52260
tagacctaga	gaggactggc	atgtggccag	ggaccaaacg	tggcacagag	agggctcagt	52320
gcaatctgcc	ccgtgggtgc	ctcccagcca	catccatttg	cccagaactg	tgacgtcaaa	52380
ccagcccggc	ccattcattc	tttattcagg	tggcataaaa	atcactacaa	aaactttaca	52440
aaagagtctt	gggagctaaa	gggtcccttc	cttgccctcag	tccccaagat	tcctggcagg	52500
ggaggacaag	agagagaaga	aggaggaaga	ctcctggcag	tggtggcatc	tccaaatacc	52560
agaggggtga	cttgggtgac	aggacacagg	ttggggacct	gaatgtcttc	agcaagggac	52620
actcttgtag	ggtaggtcag	cctccaacca	tgaagtataa	caccaaggcc	agtctaagct	52680
tgggagacca	acacttgtct	ctccttttcc	cacccagggt	gtctggaata	tgtctaaaga	52740
tggcctctcc	agcctctgct	tacaaatgtg	gagggaccct	aagttaggga	cttgccctaac	52800
ctacctctag	caaaaactgt	gtccacaagag	gccagcccac	aaaagatcac	ccccgtagcc	52860
ccttgggaag	aaatgaagat	tccccatgcc	tgccttcctc	caggccccac	cccacctgct	52920
gcaagagaac	agcttctaca	ctgggtgatgg	tccttccggg	cccaccctat	cccacaaagc	52980
tggttagaaa	gagtcacagg	agctgagagg	ctgatccagg	tggggactca	ggatgctgct	53040
gcccagggcc	cctcctcact	tgggggagct	gaactggggg	tagtcttcct	ccatgcgggg	53100
tgcaagtttc	aagtcaggac	caaaggtctt	gcctccatgg	aagtcagctt	tgtcattctg	53160
gcctatgagc	ctgttgtcag	gggaatctcg	ctgttcctgg	agctggggca	gcgcgctggg	53220
gtaggggttc	ctcacactgc	ccacaaagag	gggcagccct	atggtgtcct	ccatgtgtaa	53280
gaagaggaag	ggtcgggttc	cagtgaagga	ggagagggac	attcgattca	tggctacgct	53340
ggtagctgcg	gctgcctcca	caccagcctc	gctgagctcc	atggtagact	gatgttgac	53400
gctagacacc	accagattct	gctcagagat	cccacgaagg	tctgggccct	ggaacaattc	53460
ctgcaggcct	gcccagaaca	gcagatgact	ggtcagtgct	gccccaaagg	tatgtggatc	53520
tgtctagcat	cctggctaaa	gggaacactt	gaacccagcg	gttgatttga	atctgttaga	53580
cctcagtcta	gacaacactt	ctagaaacct	tttttttttt	tttttttttt	ttttaaatca	53640
ggatctgcgc	taggtacagg	acagaaagtc	tagagggcca	tatcaaagtc	tccatccag	53700
gaagcagggc	cacctctggc	tcaggcacac	tggcagctcc	cgtactctgc	ccagaccacc	53760
taggggcacc	ctatccccaa	gctccttacc	cagttggctg	aggggtggca	ccagggtccag	53820
ctgctgttgc	agatggagtt	taggcagcca	caccttggtg	ggcctctcct	gcagcgaggg	53880
atggtacaga	gtatcccagg	tcagggttgg	tagtacctcg	gacacgttcc	actcaaaata	53940
agtgggcatc	acgaccacaa	agctcatggt	gttcttaaag	gggaaatgag	ccacctacag	54000
ataagaaaag	gagagaacat	gaggaccaga	ccagacctgg	acctgtctgg	agtcgtggcc	54060
aaaattactt	ctgtactttt	gagacaagag	cagacaaattc	agggttagca	tgcttttact	54120
taactgggtga	agtggataaa	taccacttac	ccctttgcaa	ggtgacatgg	gaccaaataa	54180
gataatgctt	ttacacctct	ctgtgtgcac	acataagcat	atatgtttgt	atcggtgtga	54240
gtgtgtttgc	tcattgggtat	atggagtcag	aagtaggtaa	acatcagtcg	tcttcctaca	54300
ttgtctctcca	cttttttttt	tttttttttg	gtgttgccat	ctttttgttg	ttgttatttc	54360
aagacaggct	ttctctgtgt	agccctggct	gtcctggaac	tcactctgta	aatcaggctg	54420
gcctcgaaat	tgagagacc	cacctgcctc	tgccctctga	gtgctgggat	ctaagatgtg	54480
tgtaactaca	catagctccc	tccttttttg	acacagggtc	tcattggatcc	caagctggct	54540
ttgaaatgac	tgtttggggc	tggagagatg	gctcagcggc	taagaacact	gactgctctt	54600
ccaaaggtcc	tgagttcaaa	tcccagcaac	cacatgggtg	ctcacaacca	tccgtaacaa	54660
gatctgactc	cctcttctgg	agtgtctgaa	gacagctaga	gtgtacttac	atgtaataaa	54720
taaattaatc	tttttttaaaa	agagaaaaga	atgatggcta	catacttctc	tctcgtctct	54780
ctgccccaa	tgctgggatt	acagagctgt	acaacaagcc	caagtttgtt	gtgttttaga	54840
catgctaagt	tgtcccaggc	tgtcctcaga	ctctctatgt	aattcagaac	gcacctgaac	54900
ttcttttaag	gtttattttt	atcttatgtg	tatgggtatt	ttgcctgagc	atttgtctgt	54960
gtaccgtgtc	cttgacgtac	cctcacagtc	cagaggaggg	caccatttcc	ccctgaactg	55020
gttgtgagct	gcattggtgg	tgctgggaat	caaaccctgg	tcctctgcaa	gagaagccag	55080
taagtactct	taactgctga	gccacttctc	caccttgagc	ttttcttcc	cctatctcga	55140
tctaaaagta	ctagggatgg	cggatgtgcg	ttcatgtgcc	tggtttatgt	gttgctaagg	55200
gttgaaacaaa	gggctttgtg	catgccaggc	aagcaactcaa	caactgagct	acacatccc	55260
acagactttg	actcttctag	tagtagtgct	tccactacag	cctgagttct	ctatctgctg	55320
tcagcaagct	gtacaaacaa	gctatgggcc	ttcctgtcct	tgccctctcag	ttctctccgc	55380
aggtggggct	actggctttc	aaaatgaccc	atagaggagc	cacagcaaac	agtaggaagc	55440
ttggccctcg	tctttcaccc	tctcccagag	agtcagctat	aattcgagtt	tttttttcc	55500
ctctctctct	ttaaacagga	tctggttatg	tggccctaac	tatcttcaac	ttcagtcctc	55560
ctgcttcaac	cttctgagtg	ctgggattat	ggtgtaagcc	accacactca	gctcacacaa	55620
cctttttttt	tttttttttt	tttaaagaat	ccatgcagtt	aggacagcat	ggaaatgacc	55680
aggtcaggc	ctccctgggt	accagcataa	tgccctgcagg	cgggtccctc	gccagtgagg	55740
ggatggaaag	atggagccag	aggatctttc	ctctctgaac	ctcaatgtcc	cacagtgaga	55800
cactcatgtc	cactgggaga	tactgtagta	ttcaaggaag	aagcaacagg	aaggtgagag	55860
ctaagtggag	ctgagcaggc	tcgtatcctc	tcaccacggg	ctacagagaa	gtctggctgc	55920
cccctccaca	tggtcctctc	ctgcagaact	ggcaatgctg	ggcccggctt	gcccagtcac	55980
actaaccaac	agaatggatg	agcatgtgtg	gtgccacaca	cctgggaccc	cagcactcag	56040

acagctgggg	cagaagggtc	atgagtccea	agcgaacttg	tgtaacattg	tcagaccctc	56100
gaacaaacaa	aactagcccc	tcctgttatc	tcagccacag	atgatgggcc	caaggatcag	56160
tactctagcc	aaggagtcac	ggttaggcta	gaagcaagg	aagccttagc	tgagacagct	56220
tggcacggag	cttcatccaa	tcagaatggt	cagagcaata	agctttgaaa	cccgaacttc	56280
atctatgaag	caactgtgtg	gaactcctct	cttcccttac	gagcagggcc	ctggctcctc	56340
tgggctccgc	taaaacccca	gcacagagaa	cagttacctg	gcacgtgaca	aaaactcaat	56400
atattttctt	tgaggagatg	aacctcaaag	aagctgtgtc	ctggatagac	acagcataat	56460
aaacccttca	ggagctacct	accaggggac	cagactttac	ctcccagtac	caggcctcgt	56520
ttgccagcca	aaggcaaaag	ccagactgac	ctgtatctca	ggttgctcca	gcaggaacca	56580
tcgaagagga	tatgacaccg	cgtgcatcat	gtccaccgac	actgtgaacc	gctcatccag	56640
gtggaagaaa	tctttctggg	tgaggctcgg	gtcaaacttg	gtcctccaga	aacctgcagc	56700
caggcagcca	gcaggagcca	tgtaacataa	aatcagcctn	ctgcctgtct	tgcttagaac	56760
ctatnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56880
ggaaaaagga	atcttaaatt	agaagatgcc	ttgataagat	tggcatgtag	gtatgtctca	56940
ctaattgattg	atgtgaaaag	tcacgagggg	tgggtgcacc	ctgggcagat	ggcctggggg	57000
atataaaaaa	acaggctgaa	caaaccacaa	agcagtagtc	ctcaatggct	tctgctttag	57060
tttctgtctc	aggttcctac	cttgactctc	ctcagtgaa	gcagtgcaca	tgagagtgtg	57120
aagaggaaat	aaaccctttc	ctccccacat	agtttttggg	tatgatgtta	tatgtcaaca	57180
acagaaacta	taactaatat	agttgggttt	ctttttttgt	ttgtttttgt	ttgtttttgag	57240
acagggtttc	tctgtatggc	cctggctgtc	ctggaactca	ctttgtagac	caggctggcc	57300
tcgaactcag	aatccacct	gcctctgcct	ctgcctccca	agtgggtggg	ttaaaggcat	57360
gcgccaccat	tgccctggctg	gttttctttt	ttttttaata	catttataat	gcatttttaga	57420
tttaaaaaaa	aaaatggcca	tgccatataa	tataaaaaa	agtgtttaca	aatcacctag	57480
tgcccttgcc	ataaattatg	taaaaatttc	tataatggaca	tcagtctcaa	gcttaacaatc	57540
tcagcactca	tgagcctgag	gcagaggcag	gaggatgggt	agctcaaggc	cagcttagtc	57600
tacataacaa	gatcctgtcc	aaataataac	aacagtaata	atttcataca	tagaactaga	57660
aggggccact	gcaaagacag	tatgacaaaa	ccactggccc	tgccctaattg	tattttaaat	57720
aactgtcctc	ctctctgtaa	ttttcagttt	ctaattttta	cataactacc	atgtattctt	57780
tttgtaattt	taattagttt	tttaataata	gaaacaagct	aagtgtcaag	aataattttca	57840
tatgaacatt	ttcaaggcac	ttgatacata	cctcagattt	gccctccagg	tgagcagtac	57900
caattacgtg	ccaccagcaa	tgtagcttct	cttttttccc	taccatctga	ttctgtttca	57960
gtctattcgt	agttctgac	ttgttatatc	cctttttatt	gtttccctgg	gttccaacac	58020
ctcccagttg	agtgttctca	ttgaatttca	ttagcagctg	tttcattaat	ggcacagaag	58080
aaggattaca	gtgttaacta	ggatagactt	tgacaaaaga	ctatgagaac	atatcttatt	58140
atctttgcat	aaattctttt	taatcaaaag	tcctcaaaag	cctctctctg	ttcccatctc	58200
agggagtagg	tctggccact	gatgagtgtc	caggccacag	tacaggtgtg	cgtggttctg	58260
tccctgtggg	aagggcacat	ctgtgttgta	acaggattcc	tgtcttaaca	agccttgctc	58320
aggctctaag	tggtcctgag	ctagctaact	gcccttggct	ttcccttgat	taccagataa	58380
ctattcactc	ttctcatttt	gcagagcact	taccaggtag	ctatgtcctg	gaagtacgaa	58440
tgagtccttc	tattgttttt	cttttactta	aatcccat	gaaatgcgcc	agggacactt	58500
caatccaagg	tacacttttg	ctaaagaatc	actcattttt	atatgcaaaa	tgtcacctat	58560
taactgcagc	tgatatggta	catacatatt	ctctcttctc	attaaccact	aataggtgac	58620
taatgcgaaa	tattgagtaa	tttttaaaaa	tcaatactca	atttttttaga	aataatttaga	58680
gagacattca	actctgacac	cagcacccta	ctcagttcct	gagccttctc	ctgcggagg	58740
agaatctata	aataactcac	gaagctgaca	ttactcactg	tggtgcagtc	atttttttct	58800
gagaaaattt	tagcaactgt	tctaataag	cctgccagtt	atcagtagtt	gagaatgcaa	58860
gtcaactttt	aattatgcag	acgctgatta	ttcagacgac	aaattgttgg	tgccctgcag	58920
gctccttctc	gctgectacc	tttaaccgtt	ctcagtgtc	attagcacat	gttccagaag	58980
gtaggctttg	gaggggcgga	caggcactca	aaccagctaa	gcacttagag	aagctctgat	59040
gaaagatggt	aatgcagttt	gtagaattat	ctcagtaaat	tgagtcat	ggattccctg	59100
tgaattgtat	ttacatgccc	tgtccctgtc	ccccatagca	acagataata	ggattgtctg	59160
cagagagaca	acatagttct	tatatttaat	tttttctctt	gtcgaacatt	ttcacatgat	59220
ggttcgtggg	gtttcctttg	ttcattacat	ttgtatccag	actagttact	tctgataagc	59280
ggttagttag	gattcctggc	acgcgagac	tgacaccaca	gttggtctgat	cgtttcccac	59340
ttttttacaa	aaccgtttgc	ctttaagagt	cagtgttttg	cacatttcac	ccagattatt	59400
ggaaatatta	tttccctcct	gcttaaaccg	aagctgtgat	cataatttaa	gcctttctag	59460
gtagccgatc	ttcatgttat	cataccttat	tctggcatat	gtttgtctat	tacaaagacc	59520
tcgtaggtat	gcagttagaa	gcctctagtt	aaatgaaatg	ttgcgtgtgt	gatgaacctg	59580
gagtggggat	ggccttttgt	gtgccccaa	gctgttgtgt	ttcacacagt	tgttttctgc	59640
ctcctctggg	ctatcactat	cctgccactg	ccagaaaacc	ctgctgtgtg	ttcccccgct	59700
ggaggatctc	tgcttctgaa	cttctttggc	ctgagaaact	ccataaccaa	atcagtttagc	59760
attttgttta	aagagcaggt	aggctgttag	agcttgggtc	ttacatgtct	cccaggtcca	59820
cttgccagcg	ccttgaccac	tgtaactttt	tgtaaccaa	ctcatctttt	gctgcctggt	59880

ttttgggggg	tttttttggg	tttgtttaag	ccaagatcag	ttatatggcc	caggctgagc	59940
ctctcttccc	agcctctcaa	atgttagaat	tacaagcatg	catccctcag	catacctttc	60000
ctttgctttt	tttaaaatag	agtttttgcca	tagcaacaga	aatctaacct	aactaagcat	60060
agccgtgcac	atgggtatgag	gaactcacat	atgtgtgaat	ggaagttcat	agagaccggc	60120
atcactgcct	agaggccctt	ttcttccttc	cttgaggttg	tcgtgctagc	tgactgtact	60180
acaaaagagg	ttgtctgagg	cataagacta	ccttcaataa	aacatgcaca	gacagtttgc	60240
ttctctgaga	tttcagagca	gtgactacct	tcaataaaac	atggacagac	ggtttgctta	60300
cctgagactg	cagagcagtt	tccaaaaatt	ttagacaaag	ggtaggatga	agaaggctgc	60360
ggggttttgc	acacacttaa	gggtgcgtaag	taaataaaact	gagctacact	gacaggatgc	60420
tcgttctagt	agccaaccaa	agagcagttg	aaccaaagca	cctagacttc	aaacatcgtg	60480
gggagataat	cttaggagtg	ctatgcttct	gcgtcctaca	agtattatga	aactgtctag	60540
aaagcaccct	actggtaate	ccttttttgat	tatttttttt	ataaattcta	gtcctggggg	60600
tttgagtggc	acacagacat	aatggttagg	cttcgggtgtg	tgctcattca	ccttgcttcc	60660
tggggaccag	agtttgcgat	gagtcatgtt	ccatctgatt	tctgtcggat	ccggctgcag	60720
agccatgact	cagatgggct	tcaggcccag	ctgctcagtt	catcttctgg	ggaatagatg	60780
acaaggacgg	gacaaatgtc	ctgacgcaca	tttccttctg	ttcttgcact	tccagggtct	60840
aacgagagca	tcattaccaa	cagcaggcag	atacgcttg	ccacaggcat	cctccctgtt	60900
gtcagcctcc	tgaaccactc	ctgcaggccc	aacaccagtg	tgctcttcac	tgctctgtc	60960
gccaccgtcc	gagcagcaca	gaggatcgca	aaaggacagg	agattctgca	ctgctatggt	61020
gagccagcct	ttctttccac	taccctgctg	tgctcacac	ctcacatgaa	aaggataagg	61080
ggacaggaat	cagcagatat	gggcccagtg	cctctactca	tcctctgagt	ccttcttgga	61140
aagggcaatg	catccttggg	ccaataaaaa	aggcttctctg	gctgtaataa	aaaagcccgt	61200
tgagggcagt	gagccatata	cctccatgcc	ttgtagacag	cctatcctga	aaatgagcga	61260
ggagcacttt	cttggtcttct	ttcttctctg	cccagcagct	tggaacgta	ttcactttca	61320
cccgtgtttt	gttgtttttt	ctgagatgat	agggcagagt	acccaacctc	ataaggcta	61380
ggctagtgtc	tatcactgag	ccaggacccc	aaccagcac	caccatgcc	gtcacgtgat	61440
gactagggca	gcccctcggt	agagtaggca	ttgactctct	tggtgtgact	aggaactgtg	61500
ggtaatctct	ctccaggggc	tcacgagagc	cggatgggctg	ttgctgagag	gcagcagagg	61560
ctgagttctc	agtacttctt	tgactgccgc	tgtggggcct	gtcacgctga	gacactgaga	61620
gcagctgcag	ctcccagatg	ggaagccttc	tggtgtgaag	cctgcagagc	gctcatgcag	61680
gtaaatctct	gctgttccca	ggggcagggc	tccagctaaa	ggttgtcagt	cgccaggaga	61740
accattctctg	cttcccttct	tgtaactcct	cctacatgt	cgccgggtcc	gcagaaaaac	61800
acaggttgta	tttccctaata	tttccctat	aagtgacaca	aaatcttaaa	ttacacaaag	61860
ggaccaaaaa	aaaaaaaaaa	aaaaaagccc	tagaaattta	cctgctcaaa	taagtcatca	61920
aaagtgtgtc	atcaggccta	gcacttggtg	actggtaacc	ctagcactca	ggaggctgag	61980
gaagaaggat	ctcaagtcgg	aggccagctc	caagtgcac	cccatctaag	agatcaccat	62040
tccaaggagc	tatttcagag	atggtttaat	ctggggaccc	agattgtgga	ttttctgtct	62100
gttcaattcc	atctctctgt	gctggcctca	tcagacacac	tctgtagtaa	ctgtgggaaa	62160
atccgaccca	catagttttc	cctcagcctt	tgacccagag	ggaagagcca	cagtgagag	62220
catgagagca	gacccttggt	tgctactgcc	aggtaatggt	gtagacactg	gagtcttcaa	62280
cattcatgcc	ccaatgcaaa	atggtctcca	caccagagca	tgccattctc	attagaaata	62340
agtaaattgga	attggctgtg	ttgaaaattg	taaagccaag	ggtcaagaat	gaagccttcc	62400
ccagcatgtt	ttgttttgtt	ttgtgtttta	ggcagcgtct	ctctgtgtag	ccttggtctc	62460
tgccctctgc	tacctctccc	agggtgtgca	ccatgctggg	cctaagcgcc	ctgtgcatta	62520
gtgctccctc	gatcctgtc	actcttgaga	cagctctcct	tctactctgt	atcccagat	62580
aacctagagt	tcacttcaga	gcccaggctg	gcctcaaact	tgagatcctc	gtgtcccagc	62640
ttctcaaatg	cagtgatatt	tacaggccta	cacctggctt	tccttgatag	attcctagta	62700
agatgattat	cctttgagcc	atatctctct	tctgcttctt	cctctcttcc	tgagggttg	62760
atctagaatt	tattctaaag	ctgactggcc	tcagaattgc	catccttctg	cctttagnnn	62820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
tgcttaaga	gttctttgtt	tatgggaatt	tatgggaatg	ttccacagaa	cccatccagc	63000
ggagtctctg	ctgttgtttt	ttaatcttta	ttcatcttgc	gtgtgtgtgt	gtgtgtgtgt	63060
gtgtgtatgc	gcgcgtgtct	aacttgcaaa	attgcaaaat	tcagtctcct	ccttccaccc	63120
tgtaggtect	ggggatcaga	ctctgttagg	cctgggtgta	ggtgctttac	tgagccatct	63180
tacaggcccc	ccatggacaa	ccttttcttt	gaaaacctgt	ttctggcttg	ggtgtgatag	63240
ctcacacctg	tgaccctacc	accactcatg	aggaagaggt	aggaggacta	acagaattgg	63300
aagccagcct	ggactacaca	gtgagtaaa	gctatctata	tactcaccac	atggcaagac	63360
ccgtttttta	aacactgggc	aaggtgaaac	aaaagtcaat	taatttcaca	taaagtcaat	63420
agcttcatta	acggcctagt	tatcttttaa	actgtatgca	ggttagtact	tggtttcaat	63480
tttattactt	tttctctgga	acatttaaaa	gtacttttag	ggctggagag	tcagtttaaga	63540
acagtggctg	ctgccaaagg	actggagttc	actcccaagc	accaggtgg	caatcacaa	63600
tgtctgtcat	ctaattctag	gggatctgac	accctcacag	actcacaggc	agtggaaac	63660
caatgtacat	aaaataataa	ttaaaaaaat	gaaataaaat	accaggcaag	gtggcacacg	63720

cctttaaccc	cagcactcag	gaggcagagg	caggcagatt	tctgaattcg	aaggcagcct	63780
ggtctacaga	gtgagttcca	ggacagccag	ggctatacag	agaaaccctg	tctcaaaaaa	63840
aaaaaaaaaa	aaggacttta	aattgggctg	gagagatgga	ttaaaagcat	tggctgctct	63900
tcccagagg	cctgggttca	attcccagca	ctcaaattgt	ggctcacaac	tgtctatatc	63960
aacgcaatct	aacaccctct	tcaggcatgc	aggttacatg	tagacaaaac	atccatatgc	64020
ataaaataca	taagtaaattg	agtcttttaa	tgtatactag	aagctgggtg	gtgggtgcatg	64080
cctttaatcc	cagcacttgg	gaggcagagg	cagggtggatc	tctgagttgg	aggccagcct	64140
ggtctgagta	aatagagcct	tgtacttcta	cttatcacta	cagttacatt	ttataacttt	64200
gggcccttag	gcttccattt	tccactgttt	gcttaaccac	tggggcctga	agctttttgtg	64260
ctgacacttt	tgttcgctaa	tcacacaggca	accaattggtc	tctacactcc	atcacatca	64320
acacaaacaa	aacaaaacac	aacactacgg	atcctggcat	ggtggaacat	ctttagcccc	64380
agtacgtggg	cttgagttca	aggccggcct	ggtctacata	gcaagttcta	ggatagtagg	64440
gatagtcttt	aaaacaaaac	actattttat	ttatgaacaa	aacatgtaaa	gaaagaaaaa	64500
aaactgcaaa	tttatctatg	aatgaagtct	aagtaatact	tcaatattgg	aaatagcttt	64560
ctaaaatatt	tttatttaaa	gaaaactcag	caaattattc	aaacaacctt	ataaacgttc	64620
ggtataaaaag	taaagaatta	tttgcaattg	ccttaagggt	ccaagggtggc	agcctcttaa	64680
aattcagaac	aatccaagct	tcacattcca	gttcaacatt	tctacagccc	taacgtattc	64740
aaatacctcc	attctgacaa	ctgtttccccc	tcttcttttc	ttctaagctg	cttagatgtc	64800
tgtcccaggc	ttttcatgat	tttagtcatt	cacacaacta	gcaaacatta	tctagggact	64860
aaaacttgcc	agatactggg	atatcaccct	aaagggggac	tgaaagtagc	tgcaggctac	64920
agtctctaca	atctcctgaa	tgaaatacaa	agtagctaata	atttaccaaa	taaacatgta	64980
cacctgtgat	gattgctagc	tgtactagca	gaagctaaac	actaaatcta	gaaactcagt	65040
cctccaacta	gcccccttgc	cggttccagc	ctcattttta	caaacaaggg	aaagagtttg	65100
gaatggtgcc	caaagccata	cataagtga	caaaaaggag	ttggagtctc	caaatctatg	65160
gatttgggct	agttactttg	ccaaccaact	cagtaacaac	tgagctgaac	aggaacactg	65220
tggtagcaaa	agaaaactgga	actatcaatg	gcctctagag	caaaaatata	tttaaaaaga	65280
aaaaaacaaa	caaggcctgg	caaggagact	gtgagaagag	tgtgctgact	gaaattgact	65340
agttcagcca	acaaaagact	attccagggc	tggtgagatg	gctcagtggg	taagagcacc	65400
cgactgctct	tccgaaggtc	aggagttcaa	atcccagcaa	ccacatggtg	gctcacaacc	65460
atccgtaaca	agatctgact	ccctcttctg	gagtgtatga	agacagctac	agtgtactta	65520
catataatca	ataaataaat	ctttaaaaaa	aaaaaagact	attccagtgg	ggatggaaaa	65580
gttaagtgtg	gagttaaaaat	atacttcaac	tggtgatgga	ctaggtgtcc	agagtggggc	65640
aaaaggatgc	tctgtggtag	agggtcctgc	tgtgtaagcc	cagctacctg	agctcaatcc	65700
acagaatcca	cagcggagtg	ggaagagaaa	caacgtccca	gagttgtcct	ctggcatccg	65760
acgcacattc	gccatcccca	agatgtcata	catatgtgta	catactacac	actggcgcac	65820
gcgcacacac	actctttttt	aaaattcaga	cttagaggga	cataaaggat	ttgctctgat	65880
atatgttcaa	ttgaaaatga	ctttgaagat	agagggcaga	tcgaagggaag	ctcagcagga	65940
aagaattaat	aacatgcagg	tgaagggtcta	taaaactagtc	tgcaaggggc	cttggctcga	66000
caaaaaaatc	tatggggttt	gccggtaaaa	taaggaaaaa	gttgtcaaca	tgaaacacag	66060
aacactagca	agagaggagt	gttagcagaa	agaagccaac	aagctcaaac	aattaggtcg	66120
gctgaaaaat	tttaaaatgt	cttctgattt	ggctactggg	aagccactgg	tgacttcggg	66180
cagcggtttc	tctctcgtga	ccagagagat	gtctagtagc	aataatgagt	taggaggatg	66240
taaaagaagt	aaaacagccg	aaaacaagtc	caaaaagtgt	ggggtgatgg	agaaaggagg	66300
gaaacagagg	ccgcgcgaaga	tagacagcgg	catgtttatt	tgtcttggtt	tcttagatgt	66360
aaacaaacta	aaaaaactcg	tgagttcttc	tgccagtacc	gggttgccct	cagcatcttc	66420
tgatggtctt	agagaccccc	ggatgctccc	ccgcggccgt	ataatttcct	ccctgacgct	66480
ctcccgatcg	acagcggctc	cctccccggg	tctcttttgc	accgctccaa	ggccgcgctg	66540
ctagggccat	cgagcccgtc	cagggtcgtc	tctttacctc	gatggccccc	tcgctcaggt	66600
gtcccaccat	ggctgcaccg	ctaactcccc	cgctcgcgct	cttgcaaccgc	ctgagcttct	66660
ctgccggggg	cccgcggggt	gctcaacgat	tggctagagc	aactgtgcgt	gccgatccgc	66720
ccccagcgtg	agcgcgggtc	gaggggctgg	cctagacgcc	gatagccacc	gcattggcta	66780
ccgcgcggca	ggcagagcac	gtgactcttc	cgaggccggg	ttcgaggcct	agtggcggga	66840
tgccggggacg	tgaggggcgg	gcgctgggtc	gcagtgcgcc	tgtgtcagcg	cggtgctact	66900
gagttgttcc	cccgcagctc	gtcggaaact	tgcccgccca	gtcctttggc	ggacagacag	66960
aatggcaacc	caggggaacag	tcggagctct	cccctggtaa	ctgctgctaa	atatagtcaa	67020
agcagtgaac	tgggtacttc	ttcacgcagt	gcgtgcccg	cgccggtgcc	aggcccagag	67080
cttggcactg	tgggataaac	aaggtaaatc	agactcagtc	tccgccctct	tgagttccac	67140
ctgagagttg	tggccgcaag	gaaccacgcc	tcaaggatgg	tagacgcgat	atgggccaca	67200
catgtggagc	tcagagagtg	gggtcaaaaa	tcaatcaggc	tttcgagagg	cgatgcgggt	67260
tgaactgagt	taaagtgtgt	gtagaaattt	gtcaggtgga	ttccagttag	gatagttagt	67320
ttcctaaaag	cccaaattggc	ctatgcaaaa	gtattggaga	gcctggcggt	ctggctggct	67380
ctgatctgtt	tgtaatccca	gcctttggga	tgtagaagca	gcaaaaagttc	aaggtcaccc	67440
ttaacaccgt	tgagttcgag	gtcaacctga	actaaatgag	accctgaaaa	atcaaaattt	67500
gggaccagg	cgtggtggca	ttcgaggtaa	aagcaggcag	atctctgagt	tcgaagccag	67560

ccaggctaac	ataagatccg	gtctcaaaaa	aaaaaagtaa	taaaaataaa	aagggagaga	67620
ggctatatga	actgaaagaa	agacctggag	atcaaaacag	aaaactgagc	cgtctaagaa	67680
atgaaaatat	ttaaacttcat	agttgctgga	gtaagaagtc	tggaaaactt	tgggcaacta	67740
aggtaaacag	gtctagaaag	actggaatag	tagccatcta	ctgggtatttt	gatctctggt	67800
tgtacaacca	caacctacta	tagtttctca	aacagttcca	aagaatatgt	ctgggtgaat	67860
tggtagcaca	ccacagatta	actctccttc	agcatatcaa	cagctataga	aaaccccaga	67920
agaaatgatt	ttggttgctg	gtcacttggt	aggatgaaat	ctcgattttc	tagaactatg	67980
cattaataga	aagctgaatc	ttcatgttct	gactttacag	agctgcggca	gcattggatct	68040
accggtggat	gaatgggaag	cctacctact	taagaagtgg	gcttcactcc	cgaagtctgt	68100
gcaggacaca	atttctacag	cagagacttt	gagcgacatc	ttccttcctt	cttcttcctt	68160
tcttcagtaa	gtgaatggaa	acttcaggga	aatttttggtc	tggaaaatgt	tctgccttgt	68220
catttggtct	gaatatctct	tttttatagg	agagagtagc	tttatattct	ttatagtagt	68280
gggcatttag	cagttactgt	tggttttcac	gtttctccct	agtctgtgat	tactagaatg	68340
ggtaggcact	aactgctttc	ctcttttggtc	atgtgttata	cttaaggaat	gtagtatctt	68400
gctgtcgtcc	cagtgtctgc	actcatagga	tctggtgcag	gttgtgtagc	tgcccctaga	68460
agctcattga	gcctaattgg	ggagaaagaa	ccctggcact	tgggttagtg	agacccanaa	68520
cttctcaagt	tctnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	68580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnttttagtt	68640
tccaagtgcc	actttactgc	aatgtgtcac	cacatccaga	gttctgtgtt	tgtttatttg	68700
tctgtttttg	agacagggtt	tctctgtgta	gccctggctg	tcctggaact	cactctgtag	68760
accaggctac	cctcaaattc	actgagatct	gcctgcctct	gcctccagag	tgctgctgta	68820
cactaccacc	acccatccag	ctttctatat	tggttttcct	atggcggttt	aaaatagcca	68880
ttatacgtgt	gtttatcatc	taaagtcctg	gtcccaaaag	gagatgagag	gggctgttaa	68940
ggtgaaaagg	attacaaaag	cttatcaatt	ctgttcaaaa	attaaacctc	agagtgggat	69000
tagctgcttc	tttcattaga	attgctatca	gaattcactc	aggccttggt	tgcgtgtgtt	69060
attgaagaag	tctcttcctc	atcagggtcag	tgactcctta	gctcaagtac	atgcaatatg	69120
cagtattgat	aactgttctc	gcttaggaat	aaaaatagaa	ctgacttcca	cagggaaatg	69180
atgtgctgag	ctgtagcaac	gaatcttgca	caaactctgt	cagcagggac	cagctagtct	69240
ctcgcctgca	ggaccttcag	caacagggtc	gcattggccc	gaagcttctc	cgaaccggta	69300
aaccagggtg	gattggctcc	ctcgccttag	gcctcagccc	ttcccttggt	tttttggtga	69360
tcaccttgcc	ttactgagca	gtcctcaata	aatgactgag	gacttgaatt	taattattccc	69420
agcaccagcc	acaagatggc	tatgtaggcc	agttagacca	gactgtgacc	agctgttact	69480
ctggtgccct	tgaaggtctt	cctgatgggt	taagctgtgt	ctgctgcgcc	agatagttct	69540
agcagctcga	gcaccagaaa	ggctgtctga	cttccatggg	ctttgtgtgg	ctccagaggt	69600
ccaatgccat	catctgattc	ccagcttaag	gacctaagct	ccgagaaggt	tgctctgccc	69660
tcagcagcag	cagcaagtcc	tgagtgtctg	ctgggctcgt	ggtgtgactc	aggagttaga	69720
ctcggtagct	agcctgagct	gagagctgag	agaaagaaaag	gactcctctc	ttttcagaaa	69780
gggatttgca	gaactcgatg	ttagaccctg	acatggtagg	aatctgtttt	gactattcta	69840
gcctagattc	tgaagttgac	cttttagccta	gagtcaagaa	aactaatgat	tacaggagga	69900
atgtagagtt	ggttggttaa	tgttggttgg	aaaatggatg	ttagaagccc	agggtaaatg	69960
tgaggaagcc	tcacttaaca	cctcttttac	tgaagagaaa	aacataagca	accaacagct	70020
tccctggaat	gcccggctgt	tgactccgtg	agataaagag	gcattttcac	tttgacctaa	70080
ccgatagaga	ccttgcaacg	tggctctctg	gtgccaggac	tagatctgtg	tctgtgtgta	70140
ggcatttttc	ctttgaatcc	atagagcaag	ccattcagca	gttgttgctg	tgccgggagg	70200
ctgctgagag	cttcttgtea	gcagagcaca	ccgtactggg	ggaaattgaa	gatggcctgg	70260
cccaggccca	tgctacctta	ggtatgctac	cttaggtata	gccggagttc	tccttccctg	70320
ccgtgtgttc	agtgcggccc	ttgccttgct	tgtttggttc	tctcttgcca	tctgaattga	70380
cgctcttctc	cctcccattc	tgcattcctt	gccccagag	ccttaggcta	atgggtgtttc	70440
ttttccggaa	tgagacattt	ctcttctcac	aggaactgg	ctaaagtctg	ctgcccattg	70500
acagaagagt	ctccagggtg	ttgaaaactcg	ccatgggcca	tccagtgttg	aaattggcca	70560
tgagctcttc	aaactggccc	aagtcctatt	caatgggtag	gcctttcttt	ttcctagtgt	70620
ttggccaggg	cacacagtgc	tctgtgtttt	cctaggtgct	tctgtgtatg	gctttttgct	70680
acagtgtctt	aaagcatggt	gaaactcttt	tatttctctt	ttaggacaga	tatttgccct	70740
ctgcttcact	gatagacttt	aagctttgaa	ttccttccct	aggatgtgga	gaaagccatt	70800
aggtctgcat	ggagcttccc	agggaggatt	tggaggcagc	ctcaccgcgc	cttagcattc	70860
ctgtctgctt	aatcacacct	cccttggtcg	ctcactccc	tgctctctca	actccagggc	70920
tcggcccttt	ccctgggttg	cctcttattc	cttttaagac	agtgggtttc	aactagaagg	70980
gattgcaaat	ggcatttggtc	agtgtttaga	gacagttttg	attgttatgg	ctgccagcat	71040
ctagtaaagg	ctaaacctac	agtgcacagg	accgcctcca	cagtggagag	acccaagtta	71100
gctatgtgaa	ggctgagaat	ccctgctttg	gagattaaaa	aaggaagctg	agggaaccac	71160
tcagttggaa	gcacccttgg	tggcatgcac	aaggccctgg	ttctgtccct	agctctgcac	71220
aaaaaataga	atacaaggaa	gagtaaccct	aatgagctgg	tccttcaccc	agtggtgccac	71280
tgaggtcact	tgaagggaag	tctagcccca	atttagtatt	ttttgtggct	gccatacctc	71340
cagccttgat	caaactctcat	ggtatacatt	ggtaagaaaa	agggtttgaa	acatagacct	71400

gatactcggga	catggaaaca	gtatgttttg	tcagagagag	cgaaggacct	gatagacgag	71460
ggcaatatca	gagagagggc	atcagtcggg	ttagacacga	gcattccaca	gtgagcagct	71520
ctggataagc	ttttataaat	gctggttaag	gttttgaatt	tgcccaattt	tgtcaggatc	71580
ccagagtcta	tcacaaacat	acacagtttt	ctcaaactctg	ctttgcagta	tgcccgtgaa	71640
tgtctcttat	ctatactttc	agatggtaag	accctgaggg	cagaggaact	cagacccttt	71700
gtgccccctg	taagaccctg	ggggatgcag	tggacccgac	tttgtgttct	ctgcacagaa	71760
aggagtccac	tttcgttgag	actaaggaag	ggaactgaca	agcttccctt	tctggcttca	71820
ggttggcagt	gcctgaagct	ctgagtgcca	tctggaaggc	agaaaggatc	ctgttggtgc	71880
actgtggccc	tgagagttag	gaggtccggg	agctccggga	aatgaggtcc	tgcttactgg	71940
actcgtcatt	cgtccctgtg	gggcccttgg	tgtagagcaa	tcctcctcac	cctcaagaag	72000
gagctctggt	gatgactgag	atgttctgtt	ggcttgagac	tctcatcaga	gaggacggga	72060
ccttcccacc	tgacctgagc	ctagtgtctg	gcacagagag	cacttgaaaa	cagattgaga	72120
cactcacctg	ccatgctggc	tgtgtcttgc	aagagctaac	tgccctctga	tggaaccccc	72180
atgcccagaa	aagactaaat	ccagtatcta	aaggctgctt	taaagggttg	tactgagcgc	72240
cgggctttgt	ggcacacgcc	tttaatccca	gcactcggga	ggcaggcgga	tttctgagtt	72300
caagggcagc	ctgggtctaca	aagttagttc	taggacagcc	agggctacag	agaaaccctg	72360
tcttgaaaaa	ccaaaaaaat	aaaaaataaa	aataagtaaa	aaataaataa	ataaataaat	72420
aaaggggtgt	cactgatctg	caggcagctc	atgctagcct	aggcttttgg	ctcgatttca	72480
tctcactaaa	cgatgaatct	gtttccctgg	aacattccta	tggtttctag	tagtaatgaa	72540
gtgctgtgtt	ccactccagt	gagaacttca	attcttagtc	ttgtattata	attgaaaaat	72600
aataatatagc	aagaaatcag	tatgactgct	tacctcaaga	gacatacaat	tccacttaca	72660
atatactgtc	tccttaaat	tttcattaa	actgttgata	tataatttgt	gaattggagaa	72720
ataaataact	cttactgttg	gcagtttctt	cctgggatgg	caactctgta	ttgggttctt	72780
accagtgtcc	taattcttac	tcagtggctt	tcattgagtg	ttcttggcac	tactgtcca	72840
agcactgatg	caaggcaacc	ctgtagcatg	acttcatagc	acaggcctcc	ttgttagcac	72900
acctgaaagc	agaccactct	ggctgtttca	cttgagaca	gaatcttact	ctgtaagcca	72960
gtctagcctc	aaacaacatc	ctcctgcctc	agccttccaa	gttctaggtt	tataggaaaa	73020
ggccaccttg	cccagcttga	gactgcttct	tactgccatg	tctcttcagg	ctcacacatg	73080
aagtccaggg	cactccagga	ggagccgtga	gtctgtctgc	agggcactcc	agggggagcc	73140
atgactctgt	ctgcagggca	ctccaggagg	agccgtgagt	ctgtctgcag	ggcactccag	73200
gggaagccgt	gagtctgtct	gcagggcact	ccagggggag	ccatgagtct	gtctgcaggg	73260
cactccaggg	ggagccatga	gtctgtctgc	aaggcattcc	aagagcagcc	atgggcgtca	73320
ctcattggta	gactgtgagg	ctacatctcc	agatgccccg	agtgtctgtg	ttgtgagcac	73380
tgtgtctcat	ggtttccaac	tgagacagag	ggaaggactt	tgcccttttc	cctaaggatg	73440
ggtagtaata	gtccagacca	caagggacag	atagctatgg	ggttttctga	ctcactctta	73500
gtacattatt	gctgatgacc	agtttgtttg	gatgagttag	tgggaaagaa	gacccaagtc	73560
catacactct	gctttttaga	acttgtctcat	cctagccatg	cccaaggagc	agccgttgac	73620
tgtcatggca	ttacagttag	gaaataaaca	gtcctgaagg	tgccctggcag	cagcttttca	73680
agaagctggt	gttaaaagac	agtattcaaa	catctgcgga	ctgggaactg	ggcagcattt	73740
gagtctcctg	ctgtctgtta	atttaccctg	acaaggaggt	gacttgaaag	gtttgttttg	73800
tttggggtag	agcttttttca	ggaaaaaagt	ttagtccctac	agacaactct	atagttattc	73860
tagtccaaac	tcatgccttg	tgttttattc	ctaaaagccc	tgtcacactt	tgtaaaatag	73920
gtgctcttcc	tcaaaggata	tatttaacgt	tttatatatc	aggccttatt	ctgtgcatgg	73980
aagctttttt	tagatgcttt	gtaagatggc	tcagtggtta	agagcatgta	ctgtctcttct	74040
ggaagtccctg	ggtttgattc	tcagcagcta	acaccagctg	ttattccagt	tcctgggac	74100
tgatgccctc	ttctggccta	tgtgagcact	gcattgtcgt	agtgcacaga	caaatgcagg	74160
caaagcactc	atacataaaa	ctaaattcaa	aaaactcttt	cattgtctca	tgtgacctag	74220
cttgagaata	cctgtgctta	tattataatc	tagtatgagc	cagccacggg	agcaacacac	74280
ctattatctc	agactcaga	agattgagac	tagatgggtc	agagctagag	tctgggttac	74340
aaaacacctg	tctcaaaagt	aaaagggtctg	aaaaagtgtc	tcagcagcta	agagcacaca	74400
ctgcttctcc	agagggcctc	atttcagttc	ctaataccca	caccgagtga	ctcaaccacc	74460
tgtaactcca	ggtccatgag	atccaacacc	tctggctcgtc	tgcataagct	cctacactca	74520
attatacaga	gagagagaga	gagagagaga	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74640
nnnnnnnnnn	ntctagagtg	tttcaggttt	ttttgttttt	ttttgttttt	gagacaaggt	74700
ctctctatta	tgtgccttg	aactttctat	gtagaccagg	ctggactcaa	acttatagtg	74760
atccactact	tgtgcctctc	agtactggta	ttgaaggcat	gtgtcaccac	acccctactac	74820
ttcaagatct	tagatttcca	aagaagccgt	agcctagaaa	aggttaataa	gtactgattt	74880
aaaacagaaa	gaaatcaggt	acacttagag	ctgtagaatg	tcagcatgtg	acatttgtga	74940
caagttgtca	aaactttgct	cttaattcta	aagagagaag	ctgtcaaaa	acttgaactg	75000
gggctgtagc	caacttggtc	gagcccttgc	atgaagctgt	gtgtttactc	cccagcactg	75060
tggggtttga	attgatttga	accagtaga	ttcgtatatt	tgaatgttta	cctcatgggg	75120
aatgacatat	tacaaggtgt	ggccttgttg	gaggaattgt	caatttgggg	gtgagctttg	75180
aggtctctct	gctcaagctc	tgcccagggt	agaaagggag	cctcctcctg	gctgtctaca	75240

gaggacatag	tctcctggct	gccttcagat	caagatgtag	aactcctggc	tcctccagca	75300
ccaagtctgc	ctgcacaatg	ccatgcttcc	taccatgatg	ataatgaact	gaacctctga	75360
aactgtaagc	cagccccaat	taaagtgttg	tctttataag	agttgccttg	gtcatggtgt	75420
ctcttcataa	caataaaagc	ctaactaaaa	cacattcctg	ctgggcagtg	gtggtgcacg	75480
cctttaatcc	cagcacttgg	gaggcagagg	caggaggatt	tctgagttcg	aggccagcct	75540
ggtctacaga	gtgagttcca	gaacagccag	ggctacacag	agaaaccctg	tctcaaaaaa	75600
aaaacaaaaa	caaacaagca	aacaaatgcc	agcattttggg	aggtagagtt	aagaagattg	75660
ggagtacaaa	gtcgtctcag	ctagtatgtt	tgaggccagc	atggaccaca	tgagacgttc	75720
tcaaaacgaa	agaaacgaat	gaatagataa	acatttgagt	gtccagtttt	ttcctttctt	75780
tcttgctttg	tttttggcgg	tgctgaggat	taaaccagag	accttggtca	tactaggcaa	75840
gcatttctca	ctgaggaaca	ccctggcgag	tgctagtct	gtctgtctgc	ctgcctgcct	75900
gcctgcctgc	ctgcttggtt	tgtgtatgag	tggtaacctg	catgtctgtc	tgtataccac	75960
agacatgcct	ggtatctgca	gaggccagaa	gaggatgttg	gatcgccctg	aactgggatt	76020
acaaatggtt	gtaagctgcc	atgtaggtat	tcagaattga	acctggtgct	ctgaaagagc	76080
agccagtgtc	cttggtgttg	gttttatttg	gggcaggagg	tagttatttg	gttggttggt	76140
tggttggttg	gttggttggt	tttcttgaga	cagggtttct	ctatgtagcc	ttcgtgtgct	76200
tggaacttgc	tctgtaggct	caaactcaga	gatctgcctg	ccctgcctc	ccgagtgtcg	76260
ggataaagtc	atgtgccacc	aactccagac	aagcagccag	tactcttaac	cactgagcca	76320
tcattccagc	ccttctttgt	gttttgagat	ggtcacaaag	tacaactcag	actgagctct	76380
tgatcaccct	ccctcagcct	cctgactgct	gggggttaca	ggtgtgtcac	tgctcctaat	76440
tctgagtgtc	agatcttgaa	aacctattct	cgtgaccttg	atccttaaaa	caaaccctgg	76500
gagaatgagt	tctgataact	atttctcact	cctcttcaa	aaaaggaaag	ccagagaaag	76560
gggggggggc	aagccccaga	aacattgata	acttgcccaa	agttacacag	caaaattcag	76620
acagcctgca	catctcagtg	gccatctgtg	ccatatccac	cctgcccttc	tctgacctcc	76680
ccacctccat	ccctacagac	cttgacgttg	agatcagagt	ccaagccgta	tcgtaagatg	76740
gccttaggat	ctgacatcat	ggggactctc	acggtcctgt	cctcgtccaa	atgaaaatcc	76800
tgaggaggctg	tctttctcga	gtcaaacttg	gttaccact	gccctggaag	gaaacagatg	76860
gagcatcctg	agccactgtc	cccagaaagg	ccacaggtcc	acgctgtgcg	tccactggcc	76920
aaagcaacct	gagctgtcag	cagcaagaac	acaggagccg	ctgggtccca	gcatgtgtgg	76980
cacagaccat	aggtcccatg	caccacgggt	tctggctatc	ctcctgtagt	aaactcacaa	77040
ataagtgggt	gttctctctc	tgacttggat	caccacgctc	cttctgttta	aagtggcctt	77100
taatagtctg	gtgtgtggta	cgtgcctgct	ctcctgtccc	ctggggactt	ggagttagaa	77160
gccccagggc	tttctcttaa	gttaggatcc	actcttgcta	ctactccata	agatggtcac	77220
aaagcaacgt	aaaatggaaa	ttaatcaaac	cattcctgcc	acaagaataa	aacagatctc	77280
aggggaggcc	tgtggaaggg	tctcctgagg	ccttaccact	gtctagaagg	aagttgacag	77340
cagttcttga	gcagggggtg	gactccagga	gttgggggct	gctctgagag	caggacagca	77400
tgtattgtag	agtgtctggg	agggagctgt	gttatcctta	ccgtgaagat	gaggacacgg	77460
gctcatgggg	gcagagccag	gattaaacct	ggtctgattc	aaaaagccag	agatctgtgc	77520
ccagccccac	gcagccattt	cactggtcaa	ctaattcaga	aacacttggt	ctgatatgct	77580
catatgctac	aagcactgtg	gccttcagat	ctccctctgg	cctggtacct	gcattcaggt	77640
tcaccaccat	caccacacac	acacacacac	acacacacac	acactcggcc	agagacaagt	77700
ggggaagccc	tcacccttga	agtaagccac	gccaaggaga	aggatgctga	gggcactggg	77760
catcttccct	gtggaccggg	caatcttccc	tttcatctgg	gcctgcaccc	agttgttaat	77820
ctcctgaagg	tctactcgag	ggttgccctg	gaggatccgg	ggcctggtcc	cataggactt	77880
ctccagaggg	gcaacaaagc	tggaatttgac	tcgaagttct	tgagaggaaa	cagatcaaag	77940
atgagagctg	aatcagcacc	ctcactttga	aagcatgcca	gacccagct	tctgtctcag	78000
catcttccct	tgacttgctg	gggcatctgc	cggcttgccc	agaccctggc	tagggaacag	78060
tggaattccac	cgtttgcat	ccccgtccca	ggccctcctg	ctgtctccca	gagccactt	78120
cctctttctg	ttcctctgtg	gtctcactgg	ctcttctcctg	cccaccagt	ccaggcctcg	78180
cctgagcaca	cacagcctat	tgtttagaca	tcattggaagc	atacagacaa	cccaggccaa	78240
tgaagcaact	tcacgccagg	cataatgggg	cgtgcctgcc	cttcagaagc	agaggcagct	78300
ttatgagttg	ggggaccagc	tgagactcta	tagactttga	gaggtggggg	gggggtgggg	78360
ctactgactc	ctctcaaaca	caattctgga	agcactcttg	aggttcttct	caggggcagt	78420
aacagaggca	aggagctcct	tgtaggtgct	gtggatgtca	gggttggtga	tcaggtcgta	78480
gtagagagcc	cggtgaatga	cagactctgt	tcgatgttca	gctcctgcca	gagagaaaag	78540
gatgccaaag	ttcataactg	cccgtaggc	ccgatcagga	atagggacgt	tagacatcaa	78600
tccttttctg	ctctgagagc	ccgaggagcc	gatgttttag	gatgataaga	ctggacaaga	78660
tcttcagggc	gtggaagaa	ataatgaccg	ccttgctagg	aagagctcta	agacagggca	78720
aggttatcag	agctacagag	agaagagtgg	gatgtgggtc	tgaagttctc	ccatcgtaac	78780
ctaccctgtt	ctgaggagga	gccagctctg	ctcacggcag	ctgtaccct	agaacctggt	78840
taaattgacta	aaacacgata	ggaggccact	taaggaacca	aggtcgagt	ccacttacaa	78900
agtggtaggg	attgtgtgtg	tgccccccac	cgcccccttc	ctgttctctc	gacggcgcca	78960
gcatggaaac	tctgagtggg	ggaaattcag	gtccacctgc	agccttcttc	agttgacact	79020
caccagaga	aagggcagag	agggccgtgg	ccacgctgag	tggagacagc	aggacgttgc	79080

ccgttgggct	ggcactggat	ctcaggcggt	acagatcgta	gccgaagtgt	gagacagctg	79140
ctgccagctt	gttcacaggg	accttgaaga	aggggtcctc	ctcctccacg	ggctcgcccg	79200
tgctgtccgg	gactggggag	ccctgggtta	gaatacaagg	accagtaggg	aggcacagtg	79260
agtacatcac	ctcctgggtg	ggttgggtct	ctagtccctg	gggccatgag	tctgaggtca	79320
gaatgagtgt	gtgctctctg	actccacaac	ctgtgtgctg	ggaggtgggg	agtgggaagg	79380
gcaacacaaa	agggcttgcc	agacctgaac	tgtggtctga	gaacctgaag	cctggcccac	79440
tttaaaataa	aacttgtagg	gctggggaga	tagcacagta	gataaagtac	cagcatgcaa	79500
gttcaaggac	ctgggttcag	tccccagagc	tgggcacggg	ggtgcatgct	tataatccca	79560
acactgggga	ggcagagatg	ggcaggtcct	ggggctcatt	ggccaatcag	cctgaactaa	79620
tcagcgtatt	ccatctcagt	gaggggtcct	gtttcagagg	gcctgaggaa	tgactctggg	79680
ttgactacta	gcctactctg	tgtctgtttc	tgtctgtctg	tctgtctgtc	tgtctgtctc	79740
caccccctct	tgtccccctc	cctctgcagg	gaacttcctc	accaccacca	acccccaaag	79800
aaaccacccc	tcagaccagt	cttccctatt	cagcttgctg	gctggctcta	gtctgcctag	79860
gttctgctgt	gacgccctcc	ctgtctttcc	tgacaagcca	tccccctctg	ctagaccgga	79920
gaggaatttg	tcgttttctg	acctgttttc	agtgtcagcc	tcttccttat	gagactttct	79980
gcttttttgt	tttgttccca	gggcttttga	tcaaagctgg	gctcttacat	acgttaggca	80040
aatgactggc	caccagctg	tacctcccg	ccctgtgtct	tttcggtttg	gaggactttt	80100
tttttagttt	tctgttttgt	ttttggtttt	gattttttgt	tgtttgtttg	tttgttttga	80160
gacaggattt	cgctatgtga	ctctagctgt	cctgggactc	actatgtaga	ccaggctggc	80220
cttagattca	gagatccacc	tgcctctgcc	tcctaagtgc	tgggattttt	agattttaat	80280
ctgtacctac	caacctcaaa	ggaagtgtcc	atggatagag	ttcagtacta	catcatgtgt	80340
gtaacatgtg	tgaggggcctg	ggcttcaccc	ccaacagaga	agggggagtt	agtaggtgag	80400
gaataagtga	caggctagt	gcaagacagt	attgtctaag	gtcactaagc	cttaagccac	80460
acttaagcc	caaatccag	gtctaataatg	ccctagtccc	ttgtccttgt	gtgacatgac	80520
cccaccctta	cttcctccgt	atagtggcag	ctcctctgga	tcctgaaagg	agagggaaga	80580
tattcttgct	tcgatgttaa	agtaaccaag	gcctagaaga	gtgaaggcca	aagcccaccc	80640
tggtaccagg	gctgcctccc	tgcactgtct	cttctgctgt	cccacctacc	caccctactg	80700
acctcagagc	tgctggggac	gttctggctg	ctgccgtgcc	cgagcagggc	tccagtccag	80760
aggagtagca	ccaggggcctg	catcccggaa	ctacaagaga	aacaagagag	caaacgactc	80820
ccctcaccca	cacctcccc	tgccactgca	cattgcacac	tgacacagga	caagatcgag	80880
gcagagttag	cccttcccc	ccctccagct	ctcagcccca	agtggacct	tgacttgagg	80940
tcttccgtcc	ctgacctgcc	cctgcacttc	tccttgagct	gtgcccccat	gttggttccct	81000
atcaggagac	ccaccttccc	atctaagctc	cagcacaggg	aagaccagc	agcaggctct	81060
tcaggcccca	agacaatgct	ctagcacaaa	cacacaccaa	ggcttttccg	tggaggcaca	81120
cgccagctc	ctttggtagg	atttggaacc	ctgtctcagg	atgggagcag	agcccaggtc	81180
atagacttac	agaacatctg	gtctggctct	gctatccacc	aatagtctct	tgaccaaagc	81240
ctatgttaaa	gacacacaca	cactttcttc	taggtaggtt	cttgtgtatg	tagcccaggc	81300
tagccttgaa	gttgacgcta	ccctacttca	tttgccctct	gagtactaca	atgccagggtg	81360
tgagccatca	tgcctgactt	gactcccttc	ttctatttca	agagcaattc	ttagttaaga	81420
gggtatgaac	cagggccaca	ctgccnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81540
nnnnntctta	tttttagctt	gtttgttttt	cttacttgag	acctggggag	gggaggtgta	81600
tgtgtcttcc	atttgcttct	tctacctaat	aaagtcttct	ggtgatgttt	gggggggggg	81660
gagggggtag	gactcaagaa	gggattctcc	cataagctgt	tccgttttgg	tagtactatg	81720
taagggaagt	acagggtggc	agagctcggc	tctgcctgac	tggcggtgct	tgaggtaaaag	81780
gtgagatggt	gcaagatttg	ggccctcagg	agttggctct	gttggccctg	taccttctgg	81840
tctgtgggta	aggatgacca	gtaggtgaga	gatgagggaa	ccagaacaga	aggtgaaagt	81900
tagtggggcg	gagccccaga	ctagtccagg	gggggtaaac	tagatgactt	tctggaaccc	81960
caaggggctc	ggagactagt	ggtgttgagg	aagacctcta	atgtgttgta	aggcctctga	82020
actcagtagc	cgaacttgat	gccagaaagc	cccaacctgc	taaaccacaag	caggagcggg	82080
acgccatccg	ttccatggct	tcacccgagg	tggccccatg	gctgcgccaa	tcaatgagca	82140
gccgagagat	aggggctgtg	acaagccagg	aaaagttaca	gcacgctgga	aagataatac	82200
aggccaggaa	gccccaggca	cagcagggtg	gaaaagctag	atcccgattc	tgccggaggg	82260
gggccccttc	gaggtcccgg	gcaccgggtg	ccaggatcag	agaaactgac	tgaaacctag	82320
ctgacctgcc	cagaccatgg	catcctgggg	actccttggt	gctggcgctt	ccttcacggc	82380
gtttcgggga	ctgcactggg	ggctgcagct	gctgccacc	ccgaaatctg	ttcgggaccg	82440
ctggatgtgg	cggaaatttg	tcgtttcgct	gatacacagc	ctactctctg	gagttagggc	82500
gctggtcggg	tgcggaactt	ggggactgac	aaagcactga	ggggcggggg	tggaaaagag	82560
ggcctggaag	actgaagtgt	gaaccttttg	gaatggaact	ggtttggtgt	gtggatgggt	82620
gggagtaccc	agtgggagaa	tggatctagg	tctgggagaa	attgacctta	gctctttgtc	82680
ttctccaggc	tgtggcagtt	tcctcaaattg	gtcaccgacc	caattaatga	tcaccacccg	82740
tgggcacggg	tcctagtagc	agtgtcagtg	ggtgagtgtg	cagaaaaggc	tgaatcgggg	82800
aaggccttgt	tggaccggga	attctaggtt	cctcccccat	ctttggaatg	gagcagatgt	82860
tgctggaggt	ttgctgtgag	gaattaaggga	cctgagaaaa	gtgggacttg	agatatctag	82920

gctgtgcatc	agctctgagc	gaggagcctc	atagtcttct	ccggtgcctt	caggttattt	82980
cgctgcagat	ggagttgata	tgctgtggaa	ccagacattg	gccaggcct	gggaccttct	83040
ctgtcaccat	ttggcggtaa	gactctgaag	ggagaggcca	ggtagtaagg	gagcatgtcc	83100
aactcaaggg	cccaacctct	ctcttcagtg	ttctgtcctc	tgacttttcc	acaaagcccc	83160
ctgaaaacct	atcctctcag	acttggattg	agttggaggg	aggttttgac	tggttagcca	83220
ctcctgggca	ctgcccgaag	agtttggttc	tccccacaaa	cctccagctg	atcataaaaa	83280
aaaaaaaaaa	aaagccagga	atgaaagcta	gggtatgcta	tgcaaatagt	gtggcttggg	83340
gtaagagaac	ctctgggtcca	gggctgtctc	tgccccctag	ataagggtca	gcagaaagg	83400
caggattgga	ggcagtccta	aaaaatgctt	gggtaataata	aagtgaataa	ataaaaaata	83460
aataaatact	aattttttaa	aagctgatac	ctggaaggat	gaggcagaga	gtagaaaaaa	83520
catgcgtggg	tgtccctagg	ataaggagct	gggacttggt	gggcacaggt	catgcaaagc	83580
ctgaaccttg	aaccttgcct	gcaggtagtg	agctgcctca	gcaccgctgt	tgtgtctggc	83640
cactatgtgg	gcttctctat	ggtatccctg	cttctggagc	tgaactccat	ctgtttgcat	83700
ctacgggaagc	tactgctgct	ctcccataag	gccccatcct	tggccttcag	agtaagcagt	83760
tggggccagcc	tggccaccct	ggtcctcttc	cgcttctctc	ctctgggatg	gatgagtctg	83820
tggttgtccc	ggcagcacta	ccagctgtct	cttctctctg	ttctgctttg	tgtggctggg	83880
ctggttaccg	tgggcagcat	aagcatctcc	acaggatccc	gaattctgac	caagtatatc	83940
ttgcagtctc	gcccctaccc	gtttatcctc	atgcacaagg	aaaccaagac	acgtgagcct	84000
gttgccagga	acacttccac	tctcagctctg	aaaggtgtgg	aagttttctc	ttctgtcagc	84060
ccccagggag	gtggggctgg	gaagaggaga	tggtagccca	ctgcatagtc	tactatgtag	84120
caaggactag	actgtatcat	cagagagaga	gagagagaga	gagagagaga	gagagagaga	84180
gagagagaga	gaacattgta	tgagatctcc	attacagtca	ggaaatcagg	agatctaaat	84240
aactttaaaa	gtcccacagt	ctttacatat	tcttaaaatt	tcaatctctt	taaaaatccc	84300
atctctttta	aaattcaaa	tctttttaca	attaaagtc	tcaactgtgg	gctccactaa	84360
aacagttttct	tccttcaaga	gggaaaatat	cagggcacag	tcacaatcaa	aagcaaaagt	84420
caatctccaa	ccgtccaatg	tctgggatac	aactcacgat	cttctgggct	cctccaaggg	84480
cttgggtcac	ttctccagcc	aggccctttg	tagcacacgc	gtcatcctct	aggctccaga	84540
tacctgtact	ccactgctgc	tgctgctctt	ggtgggtcatc	tcattggtact	ggcatctcca	84600
aaacgctgca	tgacccttct	agtccctggg	cttcaagaga	gaagactaga	gcctggcaaa	84660
gtggcacatg	ctgataatgc	tagcacttgg	gaatgacaag	cagaaggatc	agaagtcca	84720
ggccagcctg	ggctacaaga	gactctgttt	caacaaacaa	acaaacaaac	aaacaaaga	84780
agagaaagaa	aaaactggac	atgacagccg	gaacattatc	tgacattcat	aaggtcctga	84840
gttcaatgcc	aagttggcag	tgccctagttt	gataagggtc	tagccactct	ggtaatacca	84900
tggctgactg	aacaccttac	ccagcaactt	gctgatagac	tctgccttcc	agcaaaagg	84960
aggagcttcg	ctgaggagag	aacattgaac	cctattgtat	atgaataaat	tgctgtgcaa	85020
atgatttcat	cagtctcttg	tgaatgtgat	tgctttgagt	catttttctt	ggctccagtg	85080
ttatcctggg	ctgcagtgtg	gtgtggagtt	gtggaagctt	tgagttggga	gggtttcctg	85140
ttaaggtttc	tctggctctt	ttctttcctc	ccggtttttg	ttttgtttgc	ctgggtggggt	85200
tctctgggtg	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	tagaagttgg	85260
cggggggtgg	agggggctgg	agagatggct	cagcgggtta	gagcgccaac	tgctcttcca	85320
aaggtcctga	gttcaaatac	caacaaccac	atggtggctc	acaaccatcc	gtaacaaaaa	85380
aatctgatgc	cctcttctgg	agtgtctgaa	aacagctaca	gtgtgcttac	atataataaa	85440
taaaataata	ttaaaaaaa	aagaagtgtg	catggatgat	gtagtgaaga	ctggcattag	85500
atatctctgg	atccccctgc	ctctacctct	tagacactgt	gagtatggaa	gtgtaccacc	85560
gcaccaggcc	aggctagaac	attctctgat	ctacaaatac	ctagagtatt	attcctctat	85620
gatcagaaaa	cagacccagg	gggccacaga	aatgtcttag	taggtaaaaa	cacttgcttt	85680
caggcctgat	aacctgcggt	tttttggttg	ttctgggggg	cgggagaggc	tggtggtgct	85740
gctggcctgg	aattcacaga	gatccacctg	cctctgcctc	ctgagtgtca	ggtaccagga	85800
tcacaggtgt	gtgccaccac	acttggccta	actgcttag	tttgagcatc	agtactcaca	85860
tggtagctgag	gatagaatag	actctcacca	gctcttctga	cttccacatg	tgccctgcag	85920
catgggctct	ccttcccaaa	aggaaaaata	aatgtaagaa	ttaaaaaaa	aaaaaaaaag	85980
caaaccaggg	tcttgtgtga	tggctcagca	tcaaagctac	ctcccgccac	agctgaccac	86040
ctggtgataa	cttatagcct	tgttatgctc	tcctttgacc	tccacgggca	tgctgtacac	86100
gtatgtgtgc	ccacacaaac	acacaatcaa	gaaataaatg	cagccaggcg	agggtggcaca	86160
cccccttaat	cccagcactt	gggaggcaga	ggcaggtgga	attctgagtt	cgaggccaac	86220
ctggtctaca	aagtgaagtc	caggacagcc	agagctacac	agagaaacc	tgtttcgaaa	86280
taaccacaaa	aaaaccactt	taaatattat	ttttattttg	ttttgtttat	cctgggaactt	86340
ggtctgcaga	ccaggctggc	cttgaactca	cagagatcca	actgcttctg	cttcccaagc	86400
acattaaagg	atgtcccacc	actgcctggc	ttaaagattta	ttttttcttt	ctttttgttt	86460
tgttttgttt	tgttttttct	aaaaaatttt	tttaaaaaga	accatccctc	ctagcactca	86520
ggagactctg	aagtcagggc	cagccagggtc	tactgagtga	gctctagggc	agccagggct	86580
ccacaaagaa	accctctctc	aacaaacaaa	caaaagagaa	cagacccaac	cagacctgag	86640
gacacacact	tgtaatctaa	gcccttgaga	gctgagaag	ttcaaggcta	gccacaagtg	86700
tgtggtgcat	tcaagagcag	cctgggtggg	ctacagaaaa	agaaaggagg	agagagagaa	86760

tggttaatga	agatgactct	ggaaaagtga	aactcaagag	aaagcccctc	agattttgctt	86820
aagacgagtt	gaggggtggag	aaccgccaac	gcggacgagc	cagacagaga	ctgccaaaca	86880
agttcaatcg	gttcagggtac	attactttcca	aaacgccatt	gccacatcag	gatgcttcaa	86940
tcagccaaac	caacgcagcg	actattgact	tctgcatttc	agagacttcc	gtctctgtcc	87000
agggcaatgt	cacttttagct	ttccttttga	gaaaggaaaa	gtccctgcct	ctgatgtggt	87060
agatcctcac	acacctttctg	ccagatccag	acactgggtat	gactcagcct	cggggagctc	87120
tatctacaga	gataagggtta	caaggcgtgt	gtgttttaaag	tatgtgttta	aaagtacaaa	87180
gtgagagtcc	ctggaaagggt	ctccctgccc	tcaccatcac	cgaaagcaca	aaccttaggg	87240
taatatctga	cattccttga	aatgtatgta	tgtattcatt	atgtagccct	gactgtcctg	87300
gaatggggta	taaaccagga	tgccttcaca	tctcagagac	ccatttgcct	ctgcctccca	87360
agaactaaga	ttagaggcat	gcactacat	acttggctca	tgatttactt	aactttatatt	87420
tatgttcacg	aatgttagcc	tgcattgtatg	tgtgtgcacc	atgtgcatgc	ctgggtgccc	87480
agaggccaga	agaagggtgtt	ggttggattt	cctggagatg	aagtcccaaa	caactgtaag	87540
cagtccaatg	tgtgtgctgg	agatgaaact	tggttcatcc	acaagagcag	tatgtgctct	87600
taactgtgga	ggcatatctc	cagcctcaga	tttcccagtt	aatgtttgct	ttcgcaccca	87660
ggcccatctg	cgcattgcgt	ggagacctcc	tttaccgcct	tgagcctcat	tggcctaattg	87720
tggctgggag	acttgcatga	ccaagtgggt	acaagagaag	aataaaactgg	tgtgctatga	87780
actcacctct	tctctgtagc	cattggctga	gcatactttg	cctcaacctt	ccgcccttcc	87840
ttcccctaatt	cctaaatctt	tgcctctctc	aaatgtgctc	ctcccccgca	gtaatccagt	87900
ggctgctggg	gctctagaga	gatggggggg	gggggagcaa	cggttacagc	ttaaggcagc	87960
tgcagcagaa	cttttttggct	gtatattgag	tcttaaaaaat	tcatataaac	tttgtgttct	88020
gtttctaaat	ataaccccat	ctgtttcaac	acaaaatgca	acaacaaaat	gtttcaaat	88080
gctattttga	ataattaaaa	aatttcaata	cttgatttaa	aaatgcttta	actttttaa	88140
taaattttaa	atgtttattat	ttttaaaaaag	ttacaagttt	aaaaaaaaga	aagatagaaa	88200
tcacataatg	aaattaacca	tacgcaagtg	aggctcgggtg	cactgggtaca	cagttacagt	88260
agccatttgg	agtggaggcc	atggcgcttc	acattgaatt	ttatactttc	tttatagaat	88320
attttttatg	cacctatcta	ctactgataa	caaaacaccc	atgagagagt	tagaattaga	88380
catcaattag	ctttgatcct	ctgtcataac	tctgttccac	tccctgcctt	agtcctacct	88440
catccctgtc	ctcttttcta	catcttatac	tgaatccaca	cactcagttg	tttacacaaa	88500
cacatacatc	actgtccann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88620
tgcttggtc	tggtttgctg	ttgttggttt	tggttggttg	tggttggttg	tttggttttg	88680
tttggtttga	ttttgttttt	tttttagagaa	tcttactatg	tagctcaggc	tgctccttgaa	88740
ctcacagaga	tctctgtct	ctgccttcca	agtgtctgaga	ttaaagggtat	acaccacctt	88800
acctggcccc	tttcatctat	ctatctatct	atctatctat	ctatctatct	atctatctat	88860
tatctatcta	tctaaaaattt	atctgtgtgt	gtctgtgtgt	acattcccca	gagcctgtgt	88920
ggtagtcaat	aagtaaccct	cagaagttgg	ctttctctaa	tcttggtatc	aaacttgaat	88980
tgtaggcctt	ggtagcaagc	atgtttaccc	actgagccat	ttatgacccc	atggcccagc	89040
atcttccatg	ggttctgggg	acacaaatgt	gtactttgat	gtttacagga	caagcgctta	89100
accaaccaag	tcattttccc	agccccatcc	tgactcccat	taagtgttct	ttcccccaac	89160
ccaggaccaaa	atctagagga	gtgtccatgc	ccaacaaaca	ctctgccaag	cctctcccct	89220
tactgtctct	ctcccttccc	ttccttcatt	tcttctgtcc	ttcttttctt	tctttttgaa	89280
acaggtcctt	tctctgcatc	ccaagctagc	cttgaacttg	tgatgtagct	caggctggct	89340
ttgaactcac	agctgtcctc	ctacttcagc	ttcccaaaaa	ctgggattat	agacctatgc	89400
taccacacct	ggctcatttt	tcaataaaat	aaaaagaaaa	tcaaaaagtt	cctagaacag	89460
tcacaggatt	cacaaaaact	ttggaaggag	actaaaaatg	gattttttaa	aaatgcttga	89520
agcacaaga	gttggtgaaa	gaagagagaa	gaggaaaaag	tagcttagta	ggtagaagtc	89580
aatcaagcct	cacaccctga	gttcaattcc	tgaccctatg	gtagaaggag	aagagcaatg	89640
ccggaacat	tatcctctga	cctccagacc	cgctgtggca	cgtgcatgca	cacacacaag	89700
caggagccct	tggagggaag	tcctagaaat	gaactcttact	gaagcagggt	tgccaggccc	89760
tgtgctcagc	catttttattt	ttcctttgtg	taccgcacac	gcttccattc	tcaaagttgt	89820
gagtctgaga	ggaagtactc	actgtgtccc	cagtgtgctt	ctgtcttacc	ctgggtcact	89880
tagatggggt	cacttagtgg	tagccttggt	gtggagaaaag	agaacacagg	tccgagtagc	89940
cagtagacct	gagtctttat	atctgcaaaag	gggtgtgggg	cataatcaaa	tctccccccc	90000
tccccggggt	cctgatacca	gggtgtatta	agtgtatgtg	catggtcctt	ccaagtcttg	90060
acacatcatc	caactccaag	tggctttctc	atttttctct	gccagtagcc	tcttggtgag	90120
gaaatggctg	aggaaaacag	agttgcagaa	agacagggcc	atggcctggc	tgcaggcttt	90180
ctctgagctc	gaagagggtc	agcgactctg	agaaatgaag	ctatttctga	gtgagagggg	90240
ccaaagaagg	aacacggcag	agggagagcc	cccaggagga	tggagacaga	agccggagag	90300
ggaccctgtg	cgaggctgga	ggggaggaag	aggggggagg	agttagaccc	actgtcatct	90360
gttgggcaga	gaggggctac	attcatctgc	agtatgggtg	agaggggaca	gagagtgtatg	90420
gtaacaggaa	aaatttgggg	ttgagggggg	cagcctgtag	ggctggggcc	cagcagtgta	90480
cagctaggta	gagtacacag	taactcccag	aattctctgg	ctccactaaa	tccctgttcc	90540
gctccgtgca	gagtaaaacc	cacacagggt	ggatttccagt	ctcctttgca	cccccttcca	90600

ccccccctcc	acccccagct	ctgggtcacag	ccagtcagag	ttgggggtgg	ggcagatctt	90660
gtaaaagagg	ctgggtgagga	catcggaag	tctgtaccct	ccactagcaa	agtgccagac	90720
gctccgtgac	acttttaaatg	cctcagataa	aacagtgaga	gactctcctg	gtggcaggca	90780
agatgatggg	tcagggaacct	cagcgccctc	gaggctcaga	caccaggata	aagaataaaa	90840
acaccacgga	gacccctgtg	acccctcgt	gcagagggag	aatgccaatg	tggcccagct	90900
agctgctgga	gcgcaagcct	caaggcctgt	gctagttagt	agtctactgc	tgctgccttg	90960
tcccaataaa	cccctttccg	cccaggatta	gtggacacgc	cttgctcaag	ccgagtccct	91020
gatctgccac	cttatcacac	acatacaaaa	atcccttgag	gatggttacc	atcctgggac	91080
aaagcctcat	tctctgctta	acccaagtga	cacctatatg	gcagatccct	gtgtccttct	91140
cctgatgata	acaacccttg	caatccaata	gaggggaact	cgggtttctg	tcajctcct	91200
ttatgctgat	agaaatgtac	tctgcatgtg	ggggcctgc	cttgctcacc	ctgagacccc	91260
atggggctgg	ctggggcctt	gcacatcatt	gggactcaga	gatgttgact	acatgaacgt	91320
cccacacttg	gttgacacaag	gcagaatgac	aggatgttat	gcctgggtg	tgagtgtgtg	91380
tgtgtctctg	tgtgtgtaaa	acgcctctct	ctggagccct	cctgtctgtc	tgccctcttg	91440
tcaatggctg	cacaattgtc	ctttctcttt	ccaaggacct	ctgtatgggt	gtgtccttca	91500
ttcagtgcct	ttcctctgtg	ggtttgcct	gctagccccc	tgtcactgag	aaagtcttct	91560
gtctgtcctt	gggttgcctg	gctagaacac	agacatctat	gtcttttttt	tttttttttt	91620
ttttttttaa	gatatttatt	atatgtaagt	acactgtagc	tgtcttcaga	cactccagaa	91680
gaggggagtca	gatctcgcta	ggatggttgt	gagccaccat	gtggttgctg	ggatttgaac	91740
tccagacctt	cggaagagca	gtcgggtgct	cttactcact	gagccatctc	accagccccg	91800
acatcattgt	cttgcccacg	actgctctcc	agaatggtgg	gcaggaggat	gtgaccccc	91860
acccccaggc	accgggacac	aacatcttct	acacgtgtag	gtcttggtgca	ctggcttttg	91920
tttcttcttc	caagcaggct	tcccaggaaa	tggcacttac	agagattgaa	gagtttaata	91980
catgtctcgc	tgctctctct	ttcgggaacc	ccccagagg	agcagcagaa	accaggctg	92040
gcaggggttc	tgagctgcct	gggcaaaagg	gcaggggtta	gcatggagcc	ttagccaatt	92100
tggaaaagcac	tgtgacccaa	gcacattttg	cagcagtaat	gtcaaattct	gccgttcagg	92160
catgccattg	atgtgcacgc	tgccacacag	aaaccagtga	cacaaaggca	cagccttctc	92220
caccctctcg	gtgcttagga	actaacggct	ctaattgaga	atgagagctg	aaaggagaga	92280
gacggggggc	ggccacagca	gcgcaggctg	gcactgcgtg	ttggaggagg	ctgacccact	92340
tctcgtagag	gtaagggggc	cactgaaatg	tcacttaaat	tagccaccac	tcccaacact	92400
agatctcctt	tgtccccata	cctcagcccc	acgtcttctt	ctttttttct	ctttttttct	92460
ctcctctggg	gcagcctcaa	gcccagcacc	acttttttag	agctgtaaac	caccctgggtc	92520
ctagaagccc	tcttacgtta	ggggatgaca	ggaggtagag	atcaggaagg	agggaggggag	92580
gggaggagga	aaggaaaagg	gaggggagag	agggaaagag	atcgagagag	catgcattca	92640
tcacaaaagag	ccctcttttc	tggctttttg	actgcactgt	gagttattta	gccaaacaata	92700
gatgttttat	tattttttta	gaacccgtat	ttattaacag	cctgaaagga	gagagacgga	92760
gatttatata	ggaagtgcag	tgagttaagg	ggggcaatta	agagagcaga	aagagatacg	92820
gaacacagac	ttgtaaaagg	ttttgttaaca	tccaatcaaa	ggtgcttcag	gtattttcca	92880
aggaagcaga	aggtaaaaaa	aaaaaaaaat	tgtcccata	gaagctgaca	ctggatggag	92940
caatggccca	ggcggaactc	ctgcttgaaa	gaaggtgaga	agggaggggac	acagaccagg	93000
atccgatgag	ccagagtgtg	gccatagctg	ggtcatgagg	cccagggttg	gaaggacccc	93060
actaaagtgt	gcactggcct	ttccttgaca	aaggatgcac	ctatagctag	gcgtgggtggc	93120
aagtggttgt	tattctagta	cttaggaggc	tagggcagga	ggatcaccat	gagtgtatgc	93180
ccagcctgga	ctgcatagca	acaccagtt	tcaaaaatac	aacaaaagga	agtggggggtg	93240
gggaggggcaa	catttggaat	gtaaaataat	aaaacatttt	ttttaaaaaa	agaaaggggc	93300
tagtgagtta	gttcagcggt	taagagcgct	gactgtctct	ccgaagggtc	tgagttcaaa	93360
tcccaacaac	cacatgggtg	ctcacaacca	tccataagga	gatctacgcc	ctcttctggt	93420
gtgtttaaag	tcagctacaa	tgtacttaca	tataataata	aataaattct	ggagtgaggg	93480
ggccagagca	agtagaggtc	ctgagtttaa	ttcccagcaa	ccacatgatg	gctcacaacc	93540
atctgtacaa	ttacagtgc	ctcatataca	taaaataaat	aaataaatct	ttaaaaaaag	93600
aagaaagag	gtggggcagg	ggagggaaga	agaagaaagg	taagaagcta	aataaaaggc	93660
acagagatga	gcttcatgtg	gaaacacagg	cctgtagtcc	tggcactcag	gtgggggttg	93720
ggggggctac	agtgaagta	tcatgagttc	aaggtcaact	tgggtgagac	cttgctctca	93780
aaaatacata	ngcnaaaaaa	aaaaaaaaaa	acatagccag	gcatgatggt	atacatttat	93840
agtcacagca	cttagaggac	tgaggcaggg	cagaaagaaa	aggaattcaa	gatcaggctg	93900
agctgtatgc	agtcctgatc	ctatcccttc	ccccccccc	ccagagacag	acagacagac	93960
agacagacag	agagaaaacac	aaagaaaggg	gccttcagat	ggctcagcaa	ttaaaggcgc	94020
ttgctattca	gaccccatga	cctgagctca	aagcctggga	cccaaggtag	aaggcaagag	94080
ccaactccac	agagctgttc	tatgatctct	atatgaatgc	tggggcatgt	gcctacacta	94140
tgttgtgcac	acatgcacag	attagaaaaa	gaggaggaag	aaaaacataa	gattgtttca	94200
agaaaagaaa	ggctggcttc	ttccacgtca	gtgtgagagg	aggggtctggc	ccctttgtag	94260
ccaggtcctt	cccagtcacg	tgggggctga	actgaggcag	cggaggaggc	aataacggag	94320
ctttcccaac	gcagtgtcca	gcaaaactcaa	ctctacagcc	tgtcctgatc	cacagagaag	94380
ccttcctggc	tccctcacca	atgcggggggc	attggctccc	aggctcctgg	gccccccccc	94440

acacctgtgg	agtgtctaggt	gatttgctaa	tgttgggcaa	catttgccca	cgtgggggttc	94500
ttggctcttt	ggtaatatagac	atgcctagca	ggagggcgga	gcttggagg	gggagtcctg	94560
gggttgcctg	tggctccctg	cagctggggt	gtctggccag	ctgaagaagg	agccatggca	94620
cgcaaatggg	agagcatgga	acagaggctg	tggatgctaa	gcaatatggg	aggcagtcta	94680
agcttggaag	cagcaggtgt	ctgggaacgg	gcctgtggcc	caggcagatt	tccagtgagc	94740
actccagttt	tttggcacaa	ggaacaagct	ggctgagccc	aagaggcaag	tgggtgataat	94800
gaaacccgca	gttgagggaac	agcgggtaag	ggtgccatgg	gagcccatgt	gctcatgaag	94860
aggctgggggt	gtgaagaaga	gcccattgca	ggaagccaca	catccccctg	agttccaggc	94920
agaggcagag	tccctgagtg	gggtccctg	ggtctcccct	tacctaacca	gtctcccggc	94980
accccgagca	acaaaatccc	atccaataatt	tgaggtttat	agagacctca	aaggctgagc	95040
tactgtgtgc	cactaaccat	cagcctaacc	ctccccact	gtcttctcta	gctgcccctc	95100
tttcttctga	gactgtgata	gtggcgggga	cgggttgga	gtgtgtgtga	agccctctcc	95160
gactctccaa	ccccagctga	gccccttggt	ctgcagctca	gtaacacagt	aacacaggct	95220
cagttctaca	ctggttgaga	acactcacgg	ctctctcagc	tccttagaga	gcctgttttc	95280
tcattttcct	gtccccaag	cctagacaat	ggctggtcca	tttctaagct	tatctgagga	95340
tgccaggggc	caccccatgt	ctccactagg	ctggcaatgt	tctctgtcac	tgtagtacag	95400
aagatgcct	ggtgggaggt	gagataagga	aagggatgt	ctcccctggg	gttcccacac	95460
agtgtgtgag	gtgaaatggc	agaatgggct	gggaggtaac	tctgttgcta	gagtacttgc	95520
ctagcatgtg	caaggaagg	cctgggttcc	atccccagca	ctacagaacc	caggcgtggt	95580
ggttcatgct	ggtattctca	acattcagga	ggtacagtca	ggaagagcag	aagttcaagg	95640
ccatcctcag	ctacatagct	agcttgagac	cagcctgggc	tatgtgagac	tttgtctcca	95700
acaaacaaca	acaaagcagc	agaaggccaa	ctggcaagag	gagtattacg	taaagtaaat	95760
ccatctcaaa	aagcaagtag	catgtatctt	ctttcatattt	ttttacatt	ctataaaggc	95820
ctatcaagtc	atgtatatat	gcattgtatgt	ttgtatgata	tgaagtagg	gggtctggata	95880
gatggctcag	cagttgagag	cacttgtagt	ctctttcaaa	gaacctgggt	tcaattccta	95940
gcacccacat	ggcagctcac	aactgtctgt	aattccagtc	tcaggggatc	tggcacccctc	96000
acacagatat	ccacgcacat	aaacaccaat	gcacataaaa	taaataattt	ttaaaaaaag	96060
aaattggaag	taaaactctc	taaggagaca	aaagggactg	aggggaagtg	ggaggggcat	96120
gaagggggag	ggcataggtg	tgtggtgtgt	ttaacatgca	gaatacactt	ctataaaaagc	96180
tttgggggtc	attatgcaat	gtatacatgt	gtgggtgcaa	gatgtaagct	gtgcatatgt	96240
gtggggggcca	aaggtctcct	cctcaatccc	tctctgcctt	attttcattt	aaattataat	96300
tattactatt	agtgtgtggt	gtgatgtgtg	tgggtgtggt	aagccctcac	ggcaatcaga	96360
ggatgtctgt	ggtctgagga	tgtctcttta	ccatgttcgt	gtgggttctg	tggatggaac	96420
tctggtagtc	aggtttgcaa	agctagtgtc	tttatctgcc	gagccacctt	gctggccttc	96480
aaccttattt	tttgattga	acatggaact	tcctgagttg	cctggacagc	aagtccccaa	96540
gaccctcctg	ttcctgcctc	ccccntgtcn	nnntcacang	aggacacacn	gcttantggg	96600
tntccggatt	gctgcncacc	cccccgccnc	ccnagcctcc	tgcctccccg	cccccgccc	96660
ccgctggnc	ctccccccc	cccccccccc	cccccccccc	ctccccccc	ccccnnnnnn	96720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnncagaca	cacacactca	aattaaatat	96840
agctctaact	gctgttaaat	tcacactcct	tcacatcccc	acgctaggac	tctaaggagg	96900
caccagcaag	gcccaggtcc	agcttgactt	agagcaaagc	atcctcccc	ctccacacaa	96960
tggaaacgga	cggaaagggg	catggaagca	gaaccagaca	acagcagcct	agccaagccc	97020
aggactctgc	tccttcccc	catgcctgcc	gtgcaactgg	ggaggcaaag	ccccagccgg	97080
tgtcttctga	ccgcttagcg	gaagacaagg	ggagcctgtg	attatgattt	ctgctgattt	97140
gcaatgaaac	actaatgcag	tgggcttttc	attaagccag	atttattcaa	tctaaagatt	97200
ttatttcctt	tatgtagaaa	gtgcatcttt	atatgttgtt	ggaggagcag	agatgtgata	97260
aaaagaaatt	tctcttatga	actaatagca	ctgatacata	gtggtagcta	tgcctaggcc	97320
tctctctctc	tctctctctc	tgtctcctgt	gcattgttgt	gtgtgttgtt	gtgtgttgtt	97380
atgaatgcac	acaaagtagc	ccccccccat	attatttctt	ctgtgggata	tccagactca	97440
gcaaagtgtg	gtgactggga	agtctggcca	tgcaattctt	gccttttctc	ttgccagccc	97500
aatccctttg	cattcaaacc	cgggctgctt	gctgtggcca	gccctttcac	ctggagtcct	97560
tcctctctct	tacctgtctt	cccatccttt	gcagacaatt	atcctcaata	actagccaat	97620
tacccttaag	gacaattata	ctcttccatc	agcaaacacg	ggtgttcttt	ccttgagtct	97680
tttgatgaag	tcgatattaa	agagatgctt	tatttacata	aagtcaaata	gctccctttt	97740
agaaggggtt	gggttcgatg	tcaaagtttt	aaaactctaa	ctagaggatg	ggtgtagagg	97800
gcttttggtc	agggtagaaa	agagatggag	atacttattc	tgatgttgct	ttaaaaggta	97860
ggatgcccg	agaaggtgga	aggatggggg	agggagggtc	cctcctcaag	ctaataaatc	97920
taaaagcagg	gatgagctgg	gcgctaggag	tggaaaccagt	cagaagtgtc	tgcctttgac	97980
tgaccacagc	tcctgccttc	ccctccccca	gtctctctgt	gaaccgcccag	cattaggagc	98040
taatcgcttc	agaaagccag	attggaatgt	gttgctcacc	ctccactgct	cagaaaacct	98100
ttattccagg	caaggactga	cccaaaccga	tcatggcatc	tgccaatcag	gaggccaaag	98160
gtgcgggcag	ggcgggacct	agctgtgcag	aaacagctcc	gttatggcgc	gcagaaaaag	98220
ctggggggaa	aggctaccgt	tttatctctt	ggcagatggc	ttctctcttt	gatgctttgg	98280

gccttacctg	ttactgcctg	cacttgactt	gacctaggca	aaaatagcag	cgagatacag	98340
gttctcgaag	ttagaaggaa	aaaaaaaaag	ccccaaccca	caacacaacc	cggaagtgtg	98400
ccccgcgtgt	gtttctaaag	agctgttttc	ttccaagct	ctacagcgtg	gtggctctaa	98460
tcggaatttt	ctttttaatc	atagcaggag	tcccaattag	cgtgttggtg	aatctttcaa	98520
gtagagtggg	agttccgtgg	ccacagagag	cagaggcaat	attcagcata	aagccctaga	98580
gaaagaggtg	ttgtgggcct	gtgcacacat	gtgtgtcaac	gcacatgtgg	cttgtggagg	98640
ctggcttccc	actctcaaga	tgaggtgtgt	gcaccccgag	ccttttgatt	ctcaaagctt	98700
tattaggacc	agagggactg	tgtgtgtgga	ggggtgttgc	tcacagtgc	gaaacccaaa	98760
cctggcttct	ccaggagccc	acatgccaac	aaacaggctg	cacactcttg	ctagtacatc	98820
ccctaaaggt	atggggatga	gggaccaagt	gctttgcaag	acagcaggga	cagagttctg	98880
ggagctctct	gtaccccaga	ctcagccgcc	accagggcc	agctctgac	tggcttgacc	98940
tactttcttc	tgttgttgtt	tttggagtg	ctgatgtcaa	tgcagaattc	agcagagtgg	99000
ctgagtgaga	aaaaagagga	gagggaggaa	aagggggggg	ggacgggacg	ggccgaggcc	99060
aacaggaaag	ggcaggcaac	aagacaatga	ccacaaggtc	cctgtaacta	cactaactgc	99120
ttacctttcc	tgacccccag	ggcttagcca	atatagctga	gacccagtct	tgggtgctgtg	99180
gcttcaggct	aagtaaacag	ggaagagttg	gacatgggtc	tccattctct	ctcctcatcc	99240
aacaagggga	ggaggcagtg	gccaggcagc	catgcccacc	gatgccatcc	ttctgggagg	99300
agccagccat	ttcaggcacc	tctccttccc	tgggtgccta	gaggtgctgt	gtctgcattc	99360
atctgccatg	cctgccatct	gagagaggcc	actgggactt	ggtagagagg	ttctccacac	99420
atgctggcct	ggaggaaatt	ggtcttttagg	gacactgaag	gcagtttctc	ctgttcagtg	99480
gctccttgga	aaccacagtg	acagagctcg	catgacaact	tgccggctct	caactcccat	99540
tcttagctgc	ctcaagcact	gtaaggttta	ggagagcccc	agatgtaagt	atggatggga	99600
agaccctcca	gggagtcatt	gcctaccctt	ctgaactcta	acatgggtcca	gcttttccat	99660
tccacaattg	aggagacgcc	agacctggca	ggggagcaag	cctttgtttc	tgacccattt	99720
gcaaacccca	gccactgagg	aacttgcata	caagaaactg	cctctgggcc	tctcctggac	99780
tgagccctgc	ctcccagggg	acaactgggc	aacagatcct	tccaggtggc	tgcagtgaca	99840
gatccatgct	tttatgacat	agaaaggcct	cagtctcagg	atttcacaca	ctgtatttcc	99900
ctcatcctgg	ggaccaggga	aggcgagcat	cttctgctcc	ccccaaacaa	gtgtgggaat	99960
gattaaaatc	attttttttt	tctgctccat	gaactcatac	agttttcaga	taccgaggag	100020
acaaagccct	cctgtgctga	aattagaccc	cgaaaaatag	gttagctgac	aattacttgt	100080
ttctaagtgg	agtgtgatgt	agtggcagga	gcgcaggatg	ggctgccagg	gctgcagctc	100140
ccccccccc	aaacttactg	tctcttaacc	tctcagctcc	ctgggtttct	tgccgggatg	100200
ataattctcc	ccatctccct	cctctggtgg	gctggtggaa	agcgtaatga	atcaacgctt	100260
gaagcacgct	gaagaggcca	gactcgggat	gccatgtaag	tacacagcat	cgccagccac	100320
ctctcaagtc	tacacggagc	tgattttattt	acctcccgtg	aaagagacaa	caatcatcat	100380
atttacactt	catgccgcag	cttctgcgt	ggcacggcag	cacccctccc	ctctccgctg	100440
ctgaggactc	catcaagcac	gctgccttgc	caggatgaca	gcagcccact	ctcagcctct	100500
ccctggcctc	cttacagatc	atgacctcct	gcccgtgag	gtctgtcacc	cgaaaaccac	100560
ggtacaccgg	gggctgcagc	ctctctatgg	gggaggctga	ggaaatgaat	tccgtaggta	100620
aaaggcttcc	taggaaatca	gacgctgcta	gtaattaagg	agcgaagcat	aggtgcgtga	100680
aaggtaaagt	gatgttattt	aaatgttgcg	tcatttaaag	agtgtcctgg	tgcttcagtt	100740
ccttgttacc	atgcagggtc	gtggacgggt	ggcaattagg	ctggcacggg	tagagctcac	100800
ctgctgagct	gagggagggt	ggggacacac	cttccggtaa	ttgctgctgg	gcagctctgg	100860
gtctccccac	ccccgcccc	gcccctcact	cccccccc	acttctttcc	tgacagctct	100920
ttcatttgca	ccagcttaca	gggcttggtg	cccttaccac	gaaaatcacg	ttggaagaaa	100980
tataagaaaa	agaggaaatga	aagagaaaagc	cagaaaagtt	catattaggt	tcggatctgc	101040
ggccaaacct	ggccgagaga	atccatgacg	gtccgcgcgc	atataacct	gtggcaacag	101100
ggcccgccac	aacagggcc	gccacaagag	cttcttgagt	tgccacctgc	caggagacag	101160
gatgaatgaa	tggatcatct	gtccttagag	cacaagccag	gcctgattct	ccaatattga	101220
tgtgtgaggg	agatgtcaac	agaggttccc	taaagaatga	tgcttctatt	tccatgctaa	101280
tectggggcg	tcagcttcag	tcggaacagc	cgaccgtta	ccttagctct	gctgttctcc	101340
tgtctgtgag	ccgcagaggg	aagggcgggg	tcacccagca	ttgccactcc	ccccaccctc	101400
acgtggtcca	gaccctctt	gggttgatct	gctcctgaaa	aacagtgttg	gctcaagttt	101460
gcctctgaag	gtatgtcacc	gctggctcag	ccagcttatc	tccccgggtg	tttcaagatc	101520
aaaacaccca	aacgaaagaa	aaactttgtt	tcaagagcag	agtgtggtgc	caactctgat	101580
caaagtgttt	ttcagcatga	caactcactg	cccgtgacaa	ccagtacttg	gctgttgtgg	101640
ctcagagtga	gatgcggagg	gaagtggatg	acaacagctg	tatccaggtc	caaacacagt	101700
agattcacgg	ctggcagaaa	atggctgaga	gcttggtgt	gcacccctcc	ccccctctg	101760
cctctctctc	ttttcaaggt	ggtttttgga	aatgtccttc	ctgtgggttg	tgtgcctttt	101820
ccatgtagga	cctggggcct	gtgcagatgg	ccctgtgttc	ctggtgctgc	tgttgagatg	101880
tgaacgagtg	ataggaaccc	aggcactaaa	cacacaatgt	ggttgatatc	gactagaagc	101940
aaggcaagag	caggaggcat	ttgagggtaa	aggagtgtaa	ggactgtgta	aagagatgag	102000
ggttctatct	gggaggcagg	agtcccaatg	ccagcaaata	caatggactc	tcctggctga	102060
cccaaccaga	gagaattcaa	gatggcagag	ggacaggctg	tctgagtttc	ctatggctgc	102120

accgataaat	ggtcataagc	agagtagagg	aaaaccacag	acagaaattc	atgccattga	102180
gactagaaat	ctagctcaag	gttgtgtgtg	gcagggttgg	ttcctgggtg	ttcaaccttt	102240
tcacactgtg	acatgatgct	gtggtctgca	gatgtgttgg	gctgcatcca	tagctaccct	102300
gggacacatt	catggaccgc	aggttacaca	tgctatttaa	aaactccaag	ggaagggcta	102360
gagaaatggc	ctggtagtta	agtatgcttg	ctgatcttcc	agaagacctg	agctctgttc	102420
ctagcagcca	tgttgggcag	cttacaacta	actatgactt	ctgagctcca	aagctctctt	102480
ctaatacata	catacatata	tacatacata	catacatata	tacatacata	cgtacacaca	102540
cacacacaca	cacacacaca	cacacacact	ttaaagaaaa	aaattctggg	ttggagaggt	102600
ggctcagcaa	ttaagagcac	tgactgctct	tacagaggtg	ctgagttcaa	ctctcaacca	102660
catggtggct	cacaaccatc	tgtaattggg	tctgatgccc	tcttctgggt	tgcgctgaag	102720
acagatacaa	tgtactcata	tacattaaat	aaataaataa	gaaagaaaga	aagaaagaaa	102780
gaaagaaaga	aagaaagtgt	aaacgaggaa	aattcctaata	taaaaaagaa	agaaagaaag	102840
aaagaaagaa	agaaagaaag	aaggaaagga	attctgaggg	agaatctgcc	ccttttcccta	102900
acttccaggg	ctataggcaa	cctgtggcct	ggggaagctg	tagacaacct	gtggcctggg	102960
gaagctgtag	acaacctgtg	gcctggggaa	gctgtagaca	acctgtggcc	tggggaagct	103020
gtagacaacc	tgtggcctgg	ggaagctgta	gacaacctgt	ggcctgtggc	ggcctgtggt	103080
caacgcctca	ccctctgtgc	ccaatttcc	gttctctaag	gacacatgcc	atcaaatgca	103140
taggacactc	tacatcaaga	tgatcttgtc	tcaagatgtt	taacaaaatt	acatctgcaa	103200
agacctatct	ttacatgtga	ggtcactcca	caggttctag	acatatTTTT	gaggagccac	103260
catccaactc	actatgtgac	agagtcactc	agagatttgt	gtccaggaca	gactggctgt	103320
atctgctctg	agagtccct	gcctgcccgt	gggaactccc	cagtggtcct	taagggccct	103380
gaggactttg	gatctgcaaa	gccacatctt	ccaaaaccat	tttctctctt	tggagagcta	103440
ctctaccctg	aaaccctttt	ctctgaggtg	gcttttagag	aggcaggtct	cagcaggcca	103500
ctgtgcccc	aagaagtcct	ggggagaaag	gacccaaggg	ccagtgtctg	actatcgctg	103560
agactgagaa	cattgtgtct	cacctaaaat	cgggtgtcgc	aaggaccaag	caggctctat	103620
aaatgtctta	ctgectttat	tctttttcct	ccgtctccat	ttactcctca	ttttgtttg	103680
tttgtgtgtt	tgtttgtttt	cttctgagat	gtagcccagg	ctggccttca	gctcactatg	103740
taactaagga	tgactttaaa	cttctgatcc	tttcttccct	ccacttccag	agtctctggg	103800
caggtgtgtg	ccaccgtacc	ccagctttat	ttgagactat	gattcaggct	ccatacttca	103860
tgcatattag	gtaagcatgc	taccaacttg	gctattattcc	cagcctttct	ttctttcttc	103920
tttgagacaa	tgtctttttt	tttttaatta	tatgagtaca	ctgtatctgt	tttcagacac	103980
accagaagaa	ggcattggat	cctattagag	atggttgtga	gccaccatgt	ggttgttggg	104040
atltgaactc	aggacctctg	gaagagcagt	cagtgtcttt	aaccgctgag	ccatctcgcc	104100
agtcccttag	acaatgtctt	gctatatggc	acatattggc	ctcaaaactca	gaatccttcc	104160
gcttcagcct	cctaaatact	gggattacat	gtgagccatg	gtgtttggct	tctagccttt	104220
cttccctccc	tttcccttcc	cttttccctt	ccctttccct	tttcccttcc	cttcccttcc	104280
cttcccttcc	cttcccttcc	cttcccttcc	cctcccttcc	cttcccttcc	cttcccttcc	104340
cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	104400
taatcctggc	tgtcttggaa	cttgctctgt	agaccaggct	ggcttgnnnn	nnnnnnnnnn	104460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	104520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaaga	agaagaagaa	gaagaagaag	acaacaacga	104580
cgacaacacc	ggcgccgctg	cctccactgc	catccacctg	agacaggact	caaattccaga	104640
ccaattttta	aaaccagtgt	ttcaagccgg	tacactgaag	tagtagtccc	acttgggatt	104700
atagctcctt	actttgtttt	gctttgacgt	tttgtgacat	ggtgtgatgt	agtcttggct	104760
gtcctagaac	tcaatgtgta	aattaggctg	gccttgaact	tgcttctgcc	tcctgtctgg	104820
atgatagact	gatggtgtaa	aactccactt	aggaggcaga	ggtgggagga	tcagaaattc	104880
aaagtcatcc	ttggctatgt	tgtgagtttg	aggaccaacc	ttggctacat	gatatcctat	104940
ctcaaaaaga	aataaatgta	ttgccgggca	tggtggcaca	cgcctttaat	cccagcactt	105000
gggaggcaga	ggcaggcaga	ttttctgagt	tcgaagccag	cctggtctac	agagttagtt	105060
ccaggacagc	cagggtctaca	cagagaaacc	ctgtctccaa	aaacaaaaaa	caaacaaaaa	105120
agtgaacccc	aacagtactg	ccggacagtc	tggtgtcttt	cctaagtctc	ctttcaactc	105180
tgtttaccca	ggtgtaccca	caaggtgtgt	gagcagctct	atacccagag	gtgatacggt	105240
tgtttgaatg	agagaaaagt	ttcccatcag	ctcgggtgtg	tgaatactcg	gtccccagtt	105300
ggcagtattg	gctggagagg	tgatggggag	gtgtagcctt	ccggtggaga	tgggcttttg	105360
gagtttaaa	cttccctccac	gaagaccact	gggtgatctt	tgttaccaac	agaattggat	105420
tggcctgtct	ttggctccgg	cctcagttta	tctaaaattt	acatgttacc	tgatcaaaaa	105480
ctgtttcttc	ccccaccctc	ctccctgtct	gtattccctg	ccctctagt	gtgctggctg	105540
tatactacac	tggtgatctt	gactgtattt	cagtttacct	cttgttctct	ctctgtctgac	105600
tctagagatg	tcctttcatg	gctgggacct	ggctcagaaa	tttctaagge	actgagcctt	105660
cctccatctg	aacttttagga	aacttcttgg	cttaaagggtg	tatttctgac	ttagtatgca	105720
atgagacctt	ggagcctgca	ctttgttaag	caccctgggtg	ggtggtgggtg	gtggtggtgg	105780
tggtggtggc	agtagtagaa	ctctgatgaa	cagttagtta	ttcaaggccc	atctagaaaa	105840
agaaaggctt	tgggtgcaca	tggtataaac	tcagtgccga	gagacgtgct	tcccatgaat	105900
aataatgatg	acgatgaaaa	taattctctg	tgactgttct	ccccacttcc	ctctctctca	105960

ccctagctct	tatctaccga	atccctgcac	aagcaccat	ggggtttaca	gaatctgggg	106020
cggaacgtta	gtcacttccc	ttcgctact	tcagtattgt	gtttccagaa	gtaccatttt	106080
tggctagtca	ctgaggaaaa	cggcagctgc	ctgtgggcca	ccagcccatg	ccaagtgagg	106140
tcagcaagaa	agaagctgac	agcaaagtgt	ccaactgtgg	gtctgctgga	tttctactgt	106200
gctaagtgg	ttcaagaagt	ttcttcttaa	ccccctacaa	gaaaccacaa	attttattat	106260
ctacactgtt	ttgtagatga	agaaaacacc	attccgaagc	tcactgccag	tcagcactgg	106320
aactggaatt	tgtttggcta	attcagtggt	tctcaaccag	ggtggatttg	gactccccag	106380
gggatatttg	gggacacttc	tggtcgctcat	aattgggcatc	atgtgctact	ggcacctagg	106440
gtagaggcca	gggtggtact	gaccttcccta	ctatgaacag	ggcagccatg	tacataaatt	106500
ctctcgatca	aaacatcaac	agtgctaagg	ttgagaaatc	tcaggctgaa	accctgtcat	106560
ttggccttga	ggtgggtggg	aggaggttag	agagtgggaat	aaaatcagaa	gggccaccac	106620
agaggcctcg	agtgaggagg	gaacagggct	cctatgctag	ggataatgga	gaatagggca	106680
gcttggtgaa	acttttcttt	cttccaagct	tggttagagc	cctgctcaat	ttcccccaac	106740
tctgccaaagt	cagtccccgg	acttcgcact	aagtttgtcc	tggagtgaac	ctgactccag	106800
cttggagctg	gggcaacaca	tttacctggg	tcttccagg	agtgggttaa	aagtcaaaga	106860
taagtggtat	ctgagaagtt	aaaagtgggg	ttagtgggta	aaggcaggag	gaccccatca	106920
atcagctgac	ctaggaagga	agagagcaac	tgaagcagca	aagagctggt	ccaggagact	106980
ggattctgac	ccactaggct	tatcttccac	agcctttctt	gtttaggctt	gggctcagtt	107040
tccctgcatc	atccgcagga	gccccctggag	caccacatt	cagccggccg	ccaggacagg	107100
ctccccagca	gtggcctccc	cactaactga	cagtgggtgac	aggaaatatc	tccattcca	107160
atctcctcag	aagtctgaat	aagtaaggga	catagtgttg	ggagaggcgt	cactcttggg	107220
ttgatgaaga	aaagatcatg	agaagcatac	attttaccgg	ctatggttgg	ggttccatta	107280
ccacccatga	cggtgttag	ggggaagggc	agaaaaaagg	agatggagaa	ggacagacac	107340
gtaagaagga	ggatgtggtt	agggtgatt	cgtccccctg	ggtcgaacaa	tcagctgtta	107400
ctggggcgga	aggcaggaag	tctccttaaa	gacacaatat	tctgaacgtt	gaactcagga	107460
tttgaagcaa	gcccagcagt	caccttagtg	gagcccatat	ttaatttaac	acagaagcgg	107520
ctatctcagg	cttccctctc	atcttcttgt	ctcagatgct	ctcacttaga	aagcttagat	107580
gctcttagaa	atgactcaaa	agtcaagaac	cccaggcca	aagtttctct	ttgggggtgg	107640
ggagggagt	aagaggtggt	cccagtcttg	tccctttaa	taagcaattc	agcagctttt	107700
gccaagtcat	tgggtttcatt	tcggttttgg	cccatcccc	gcctttcaga	ctctgattgg	107760
ccccctagga	aggagccgcc	tcttcatttg	tctccacctt	tgaatcact	tccctaagta	107820
ggcctgagtc	agagaagcgt	ttcgaggggc	gggactgaat	gggtgttaat	cttagaaccg	107880
ggtttctggt	tgatactact	ttggtaaaga	tcttccccta	atttttaaaa	agacgcttcc	107940
tctctaaaa	tgagggcgaa	tcctttgtta	agaacgtgcc	ccttgagaag	ccgtgggctc	108000
ttcagcgact	aagacgagac	attcactaga	aaagatttca	ctaaaccac	gagggataga	108060
ctagacctcc	agtgaagatt	gggcctgtgc	gggtgacatt	tgtccctata	ccccgaagac	108120
ctcgagctag	cttccagtg	aagactgggg	ccgtgcgagt	gacagtggtc	cctatacccc	108180
gaaaaaaaa	aagtcctatt	tgtggaaaaa	aaaaaagact	tcgggtgttc	tgctgcatcg	108240
gtggctggct	tccatcttta	gttctactca	ctcctgttgc	ttcgctgct	ccaccttcgc	108300
ttagctcagg	cctcctgtga	atcagttttg	aggctaaaag	aagttccaag	aaggaggggc	108360
tgtagccctt	taaggacttc	cccgcgaccg	agtcagagat	cagtttataa	atgccaaact	108420
acagagcgcg	ctgcattctg	ggaagctgag	tgtcacgta	agaacttcat	tgaccggaat	108480
gcactgcaaa	aatacacgcc	tatacttcc	tctgctcttt	aaactgtagt	ttgacgtaaa	108540
gctggtctaa	gcaagtcgcc	taggccgagg	gttagccaca	ccttttcagc	cattggccag	108600
ttggttagtt	ggttaggcgtg	gcttagagaa	gctcctccag	gcaaggggg	ggcctccttg	108660
ccaatcagag	cccagacgcc	tgaatgggcg	ggagtaagca	gaggtgctgg	cgcccccgag	108720
tgggtgtggt	cacgttgccc	agcaatgggc	ggtgattggc	cctgggtggt	tcattcgag	108780
ctcgtgcgtc	acgacgcgcg	cagctgatcg	gagactggag	ccggtgtgtg	ctgggcgctg	108840
ggaagagaca	gagcggtcgg	ccgtgcggac	aggtcgcagt	gattttgctc	ctctgtccac	108900
agcaaccccc	gcacccagca	tcaggtgggt	gtgatctggg	gacccgggtca	tcccgggggg	108960
aaccgcggta	accgggtgat	ggggaaaagta	gggtcctgac	ggccacaccc	tgcccttctg	109020
ggggagggga	gagggggcgg	cggggacagg	ggcgctcttg	ggagaggagc	ctggactctc	109080
ccgagtagtg	tgtctggacg	tttaaagaga	gagtcccggga	caggagtcgt	ggcagaaggt	109140
ttggagaagt	aactggggag	gaatatgaga	ggccagaggg	ccgggggcgt	ctaaccgccga	109200
cgcccttttg	tttgaggatg	cccagactga	ccatttagcc	tagggaggat	ctggagagc	109260
gaggggtgcg	aggggtgcatt	gcctctaccg	gcgctgactg	ggtcaggggc	agttcaagtc	109320
cctggcaggg	aaggggtcgc	tgggcgggtcc	ggccctcct	ctcgttccct	cccggggatg	109380
ttatgtaagg	ggggagggga	aaggagtagg	gggcggcggt	gcggaggcct	tatgcaaccc	109440
aaaggttagg	gtttcacccg	gggttgggcg	gaggttgggg	ggggcggaca	ggaggagtgc	109500
ctggaaactc	taccgcacc	ccccctccca	gcctaaactg	ctgtcttggg	cagagagaag	109560
gtcacctttg	cacctcccc	ctagtatgtc	cggtagagag	gcccctagcc	cgggcttggc	109620
ctgactgcct	gggaagccgg	ctggctgggt	ggggcgctcg	ggtagtcat	cgctgggctc	109680
cctctctccc	ccctcctgg	ccaactcttg	gccccctccc	acggcctccg	gttaggctaa	109740
cgttcccacc	tccctctggc	cctagtttca	gtctccaact	catttggcct	gtcacctctg	109800

ctgttagagt	aggctagaag	ctgtcatggt	gccagagagt	tgatggagca	gctgggtcaga	109860
gggtcagtg	cctgggccc	ccccgcccc	cagccaaggg	cacctgcttg	gcacaaactc	109920
tcagcagcca	gtgaaccctg	tggcctgaac	agagctatcc	tgggcagaga	gaagtggaca	109980
gagactgatc	acctaggaga	aggaagatcc	gacaaagttt	atacttccca	agaggctttt	110040
ggaatttgaa	agttagccac	cctagtgtaa	tctttccact	ctctgaaaat	agaaatccca	110100
aggcaaagtc	tccttgcccc	ttctatctgg	cagtggccat	gtccttgga	tgactgtgca	110160
gaaccaccct	ctcgggctcc	cagccctcta	gcctggcacg	ccccagccc	cctccctgag	110220
ccatgctgta	gggccccggc	ttttactgct	gattcatgcg	ttggaactgt	gggggcgggg	110280
cttggaaactt	ggaacaaagt	tcagacgtgg	aggggcccgc	agacagcctg	gaattcatac	110340
cagatgtacc	cggaatgtgc	aagcggaatg	cctggcatct	ctagtctga	ggaagctgcc	110400
cagccaccct	acccatacct	ccctcccttc	ctgccttttg	tcagctgtcc	tcctcagac	110460
tcctgagagc	ccctgctgac	cttccaactc	tagtgccctt	cccatttcta	accctacaca	110520
aaccctcctt	gctgctgaat	tccttaagaa	caagtcattt	gagttgatca	cagagctcat	110580
atthctgaag	tacatthttt	ttthtaactt	gggactttgg	ttctacacc	tgccctttga	110640
atgccgaaga	tgctgggctc	cttagcaggt	tgccaagctg	tgccagctcc	tagtctgtaa	110700
aggggcacaa	agcaagtgca	tttagaagcc	cttgcttctt	tattcaagaa	cccctcatta	110760
gaaggtaactg	aaagtcagct	agagccaggt	ttggatggcc	tctgggtcgc	tgccctgtc	110820
acccagcttt	cctgtttttt	tttttccctc	ccttcccttt	aggaacctgt	gcctcccaca	110880
ccctcacctg	gctgagccgc	agtagttctt	cagtggcaag	ctttatgtcc	tgaccagct	110940
aaagctgcca	gttgaagaac	tggtgccctc	tgccctggc	ttcgtggagg	aagaggagaa	111000
gcagcagctt	tgccctatcat	ccggaaggtg	acagaactgg	ggtgggaagg	tctggcagc	111060
tggggtagat	gctttatggg	agggaaaccc	tgctcctctg	gggagccctt	acccccactg	111120
gcccagtgaa	agatttaggt	taaaggcact	gtctataaat	tggggaatag	gtgactccac	111180
ctccccaaaga	ttagttgatg	tctgtgtggc	agtgggaaga	aatagaagga	aaagtctgtc	111240
tgthttactga	gacttccctt	taggcctgcc	tttcttatct	tcatcatcac	catgccaca	111300
cacacacaca	cacacacaca	cacacacaca	catacacaca	catacacaca	cacacacaca	111360
cacacacttt	cctttccatg	aggtccaaaa	gtaaatgtac	tcaggaagg	ggacattgaa	111420
actccgttct	aagtagtcat	ttgtgtatth	actthttttg	tttatttgtt	tgattgactt	111480
tcgagacagg	gtttctctgt	atagccctgg	ctgtcctgga	actcactttg	tagctaggc	111540
tgccctcgaa	ctcagaaate	tgccctgcctc	tgccctccaa	gtgctgggat	taaaggcgtg	111600
tgccaccacc	gcccggctgt	atthacattt	cctttattht	ttthagtctg	gcccagattt	111660
tgggttttag	ggtacttacc	cttacacctg	tggattthtc	cacctgtata	atggggaatc	111720
ccatagataa	gtaggcagga	gggcattaaa	agtcaccagg	tggtgactca	gagcctgggc	111780
tcttcttctt	ctcgtggatg	gaaacgaaac	agctcttcac	atgaactgtt	gtccttcccc	111840
cacccctga	ctactcacc	agctcagggg	gattagatg	gaaggaaagg	ctatgggttaa	111900
gtcccaggca	agctcgtggg	aggtcagctc	tctactggct	tctcaccatg	catgggtggt	111960
ccaaggcttt	ccctccacct	aaagcaaaac	tgtagctctt	ggttgggttc	tagcaaccac	112020
tgccatttht	tttctgcctt	tgctthccag	gatagtgaga	ctctgctcaa	tactgtgcag	112080
gcaagaaatt	gtcaggggag	atgggttgta	tgatatgagt	cccttctgct	gcctctagct	112140
cctgattcat	tctcacgtat	gggcttggtc	tctgattgtg	gttcaccttt	ggcccagtct	112200
tcctaacaga	agatgggttc	agggggtaca	ggaggctggt	tgthgtatth	gacaggagga	112260
ggagttctag	cctgttcccc	atthgtgaga	aactgaaagt	cataggggag	ctatagttcat	112320
ctaattccagc	cccactgcag	tctaagctga	gggataggat	gtgtaaggga	ctgtagcaga	112380
cgggctgggg	aggctgagtc	ggctcacaca	ttgcgacaaa	gattgccctt	ccctcgacct	112440
cgcttgctth	ctthctcctc	cccttccctg	gccacagtgt	gtccctccag	cactgggtac	112500
atggctctgc	tgctcctcatc	caacatggag	cctcagaggt	gagaaagggc	agcctggaag	112560
caacagaggc	aggcacaaga	cagtggagga	cctggcctgg	aaccacaagg	gcctatccgg	112620
acattggtca	gagaggcacg	tagaagcctg	gagaacacca	ggaaagagag	cagccagcca	112680
gcctcagtg	aagacacgtg	cttcacagcca	tctcctctca	ggacctgcct	tcctgggaga	112740
tgaaggcct	ccaggaagta	tggtcccatc	tctaccctgc	agthttctata	aacagcctca	112800
aggagcatga	gccacctctg	aaaggaaata	cacagcaaat	tcaaaaagag	attcaaatgt	112860
gtaacactgt	gggaaaacat	atctatgact	ggggttgtag	ctcagttggt	aggtthgtct	112920
aacatgcacc	aagccctggt	tctgtcttct	gcattgcata	aaactgaaca	ggttgcccca	112980
ggtctgcaat	cccggcactc	tggaggtggt	ggcaaaaggag	cctacattca	aggtaatcct	113040
ctgctataca	ataggttctg	agccagcctg	ggctatatga	gactgtctca	aaaaataaaa	113100
caaaaataaaa	taaagcattg	gttagtaatt	caaagaaagc	agatgtggct	gaaaccgttt	113160
tcctgatca	taatacaaca	agcaaatgaa	agccagaaga	aggctcctgt	gccttggtgtg	113220
tggcagtagc	aaccattgtg	agagatgcct	ttggacctgg	tagthttgctg	tcttagaaat	113280
gtatcctaaa	ataaggattt	ggttataaaa	tgthcatctc	aggggttgtaa	tagagaaaaa	113340
tggaaagcag	ctgthttgth	ggaagtccat	tcctthttctg	ctgtcatgaa	aatgtatagc	113400
tagggcttgc	ctaagtaaat	tataattcatc	tgaatgtggt	gttctgtgca	gccattccaa	113460
gtcttacaga	agaaaaattg	agtggaaata	taaatattga	atactaaaaa	gattataaaa	113520
gtatgagtht	gtgactgtht	ttaaaaatag	aacacatact	tgtaatatat	thttthtaaaa	113580
accatccaat	tgagtggaaa	tataaatact	gaatactaaa	aagattatga	aaagtatgag	113640

tttgtgactg	tttttaaaat	atgaatgcat	acttgtaata	tatttttttaa	aaaaacactg	113700
aaagtggatt	caaaatggtta	agaatgggtg	tttttgtatg	gtgggatagt	acaattgtga	113760
attttccccct	tgttttttct	gtctttctaa	tttttaata	ttgtgcattg	ctttcatatg	113820
ttaaataaaa	tacaaaagac	aaataaatgt	tttaaaat	ttactctttt	atgagtgttt	113880
tgccgtgtgag	tggcaggaat	ccaacattgt	cttctgaaag	cagctagcgc	taacttctga	113940
gccgtctctt	cactccctct	gtaattttta	aaaaatatat	ttgtatgtta	tattatgtgt	114000
ctttgtgcac	cagagtgtgg	gtgcacattc	tgcagaggcc	agaagagggc	atcagattcc	114060
ctggagctgc	acgctgtttg	gatcttctga	tgtggatgct	cagaatcgca	ctcaggctct	114120
ctagaagagc	agcaaatgct	cctagccact	aaagccatct	ctccctctag	tcctcattgt	114180
catgttttaga	ttttggagaa	tttgcttgta	ggaggatggg	ctacaccaag	tgccagggtga	114240
aagaaaatgt	ttgcttggga	tacctattgc	ttcttgagtg	tgtgtgtgca	tgttgtgtg	114300
tgtgtgtgta	tgtgtgtgtg	tacactggag	ctaggaatca	tatccagggg	ccttttcaag	114360
ctccaccaca	ctaaagtcaa	ttctatgaac	ttcattaatt	gtctgaatcc	acctactctc	114420
tacacacagg	aagcattcct	ctgactttct	gactgtcagc	cagctaagga	gggtgtggct	114480
agaataagaa	agaagggaan	tgctcaaaac	ctgtcactct	ntggggnnnn	nnnnnnnnnn	114540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114600
nnnnnnnnnn	nnnnnnnnnn	nnnnnncccc	tattttacca	agcattgcag	gataaataag	114660
aacatagctt	aaagaattct	agagaaaaaa	aaaaaagtcc	actaatgttc	aatgttttaa	114720
ttaattttta	atgggaaaaa	gtaatatcaa	agaaatccat	gtctacggca	tattaatgtg	114780
gttaatgcaa	acaggacaga	ttatcttggc	caaattaaac	cacagtaaca	cgttttgaac	114840
taaaagtcat	ctattttatac	gagatgaaat	gtgaattagc	gtgcgcgcgc	aggtgcacac	114900
atgcacacat	acaaacagag	tctttttatt	cttaagccta	ccaaataact	cttaagcag	114960
taaataagatt	taagctctaa	aattttaaaa	atagtgaac	cccttcaggc	tatgacagaa	115020
tgctgctttt	gccattcctt	ctgataaaag	tcccaaaggg	tgtcataatc	tgtatccttc	115080
ttagaaaagt	aaggagcaca	tcctatgagc	tggtagacctg	agttttacac	ccaagtcaca	115140
cgtagcacaca	cagcaatgtc	ctggaccatt	tgtgaggagc	cccacgctgg	tcctgagcaa	115200
caactcactt	ggactggtac	ggggtgaggc	tggcgatggg	caccactttg	gactgtgtcc	115260
cacctgaagg	ctgcagcagg	ccagttccgg	caggttttcc	aaatggcttt	gaggcaccat	115320
aagccttagc	tgagtgga	cctatgcaaa	cacacaggaa	aagggcagtc	attcccgttt	115380
tatttctcga	agtgaatgta	caggcaccat	ggcacgcatac	tggagggtcag	atgataatct	115440
caaaaaaatc	caacttttgt	gagacaggct	cccttcattg	tcttcccacc	acacttgcca	115500
ggtagggggc	acaagctccc	aaggagtatc	tgtctctcta	agggcactgc	catcagagac	115560
actctgacta	ctacacctgg	cttttaggtg	ggttctgagg	atttgaacgc	agatcatcat	115620
ggttgatag	caaattttta	ttcaccaagc	cagacaaccg	ttttttaact	ttttttaaaa	115680
aaagatttat	ttttatgagt	acagtgtggc	tgtcttcaga	cacaccagaa	gaagacatca	115740
gatccatta	catagtgttg	tgagccacca	tgtggttgct	gggaattgaa	ctcaggacct	115800
ttggaagagc	agtcagtgtc	cttaaccgct	ggccatctc	tccagccctt	attttaacttt	115860
taaaagttta	aaattaattc	tcattatgca	cgtgccacag	ggaagtgtaa	cataagcaca	115920
tgtgtttgaa	gtatatatgg	ggatgtagtt	atgtgacggc	cagaggacaa	ctctgtgaag	115980
ttgattcttt	cctttcactg	tgggtcctcg	gagcgtcaat	agggtcagca	ggcattttac	116040
cagctgagct	atcttgtggg	ccaccaaaact	tattttttaa	aagctagaag	ttggttaaag	116100
agaaggaatt	catgtaattg	aaatagaaaa	gcagagtgtg	agaggaggga	gaggagggtg	116160
gtatgatatg	tgaagaaaaa	ccaagaaaaa	acctctacaa	gtaaatctag	tcgtgttcca	116220
cagatcacta	tgttcaagga	tgggtgtcca	gaaaagggtc	ctaaaacaaa	gccagggaag	116280
taatcaaaat	gatcccatcc	agacctgcta	cagagaagct	ggggggaagg	agggagaggc	116340
cctgtaacct	caagtccccg	ggcagcacag	cagagctgtg	cccaatgata	tcctgtgatac	116400
aggacaacag	gatatctgag	aaggagtccc	agttagcgtg	tagcagaagc	cagaggcctc	116460
gaacaaaacc	aacaacttca	ccatgatgaa	catcataag	atgtgtggag	aaaagggtgc	116520
actctgggac	acagtgtcac	acactacagc	ttccactgtg	gacaaaaagt	cacacactac	116580
agcttccacc	tgggacacaa	tgtcacacac	tacagcttcc	acagtcagat	tattttttctt	116640
ctagcggaga	ggctgcaagg	gtggaggggc	ggtacaaggg	aacaggggtg	agtgggatcg	116700
gggtgcatga	tgtgaaactt	acaaagaacc	aataagttag	gaaagagaga	tgaagacata	116760
gaagccatgg	ccaacaaccc	cagcactctg	gaggcaggca	ggagggtctc	catgccaaag	116820
gtaatgaatg	ctatgtctaa	agagatgata	aagtcccatg	agagttacaa	agtcagagga	116880
gaaacttgga	cagaaaagcc	aaatgtaaca	atgcatgtgt	aggtggggaa	aggtggggat	116940
ttaaaagaaa	tttcacaaaa	aggcacagaa	tgtaaaggtg	aagcgagcga	ctaccagata	117000
aagcatcaac	taagaggagt	caggaaacagt	gtatttttaac	tacctttata	aaaaaaatac	117060
tcgggctggt	gagatggctc	agcggttaag	agcaccgact	gctcttccaa	tggtcctgag	117120
ttcaaactcc	aacaaccaca	tgggtggctca	caaccatcca	taacaaaatc	tgatacggtc	117180
ttctggagtg	tctgaaggca	gctacaagtg	tacttacata	taataagtaa	ataaataaat	117240
aagtattaaa	aaatatgcaa	taattggaat	tattctttgc	aattctacta	tgaattggagt	117300
ttttgttctc	gttgtttgtt	ttccgagcct	gtttcatgta	atccaaactg	gcctgggaacc	117360
tgctaataatg	tggaagatga	ccttgaactt	ctgattcccc	tgtctccacc	tcccaagtgc	117420
tggcctaggc	tactctgcct	gactcaggct	cagtgaattt	tcaaacactg	cttcaacctt	117480

gaccctaac	ttccatctct	tgggtcccat	cttaccatt	cccagactcc	cattctgtgg	117540
ctggggcttg	ctggctggtg	gtgtcgtgc	agaggctggg	gaagggaactg	cttgctgctg	117600
ctgctgctgc	tgctgctgct	gctgctgccc	atatacctgga	gatcaaaaca	ggccaattca	117660
gtgcaaacc	agtcaagaga	tttcaacacc	agcagtaa	accttaggaa	aaccacctt	117720
tggctcttag	aggagcagac	tgcaccacgc	tgcacaccct	gtgtctgcaa	agggtcact	117780
tttttgtctg	cgcatagtaa	agactggttc	catttccctt	tcaactgttt	agaattgaca	117840
gctttcagtt	ttaatgaggt	tctccctgaa	gccttgccca	tttctccttc	aagaacctgc	117900
agtaaaagtca	ctgggctcaa	tctgtgatac	taccactatt	tccatagcaa	caaactcgatg	117960
tcatcaacag	gacggtgtgt	tctgtggcagg	atccattaga	agagaaagca	gggtggtaag	118020
gaaaacctaa	aagcagttca	attgtctctc	aagtgtttg	gcttcaaagg	aaaaggagac	118080
gtttacaaag	ggttctaccc	tttcccccaa	acaaagcaac	cttatttttgc	aactgacagc	118140
tgggtgaaaa	tatgagttga	aagagtctca	ggtagcgtga	agtgtctagt	aactgtcaca	118200
gcaggagag	atgagaacag	agttcagaag	cagccgtgtc	aaggcaacag	ggagataaaa	118260
aggaggggga	ccgagcacgc	ttctaaagca	gcaaataaga	cgcgccacc	tgctggagtg	118320
tgtattttg	cttcaccctt	tatacaaatg	tttaccagg	tgaagaatgt	taacattgta	118380
aattgtctcc	cactatttct	aaaataacga	gtttatggg	tttgatttta	cttctgtgat	118440
tggctctgaa	ggctgaagcc	agcctcgttt	aggctgtctc	cctggagtg	ttgtagatga	118500
caaaggcagc	taaaaaaaaa	aaaatgtccc	agagctcctg	aaacactaaa	actggtgcgc	118560
acaggcagga	agtctccctt	gccactgagg	ctctcccttc	tccaactgta	agttctaa	118620
cctgctggg	ctcctgaggt	cccaactcac	cgggcatgct	acatgcccat	ggaattcatt	118680
tttggcaaaa	ttactcttaa	gtagctcaag	gaatgagacg	taattgtgtt	gctggagaca	118740
aatacattca	gcacaaagtt	gagaagatta	aatagaattg	attatgcttt	agttcatcct	118800
acagagagaa	aaagtgtgac	acctatttac	ataggagagg	ggcccaggct	actgacatg	118860
agcccttct	ttacagctca	gccttcccat	tatcactcaa	ctcagacacc	cagccagtt	118920
cagtctatgt	gaagcatttt	taaaatcagc	aagagaaagg	tgagttgctc	agtagtacac	118980
tgaaggaaag	tagaaatgag	cacacagtga	cctgctgcat	aagacacagt	taaaagggtg	119040
acctatcttc	cagcgaagtt	cttgcccttct	ttaaaaagaa	tgtggtattt	gctggtatgt	119100
gcacctcacg	aagggtgcttg	cagctatgtg	gagggcagag	accatcctga	gggaatcggg	119160
tctccttcca	ccagatcct	agggatggca	caaaggcac	gggcttgaca	gatgccttta	119220
cctgtgagc	ctttcccca	gcctctcctt	tctacagtc	tccaaattac	ctgaggtagg	119280
agcctcatt	tcagaacccat	ttctgctg	catgtctgac	tgaactggat	atggtcactc	119340
gactcaattt	ctataaaaag	ataactgagg	agccggcttg	gttggtagag	aaaacataaa	119400
gaccogagtt	ttgttcccac	atggggcgtg	cgtgctagag	aggtgaggac	ggactaggca	119460
gggagcagag	gtaggcagat	cccaggcctc	aatagccagg	cagcccaacc	tgagctctaa	119520
gttcagtaag	ggagcctgta	accacagaag	acactgtgcc	gtgtaaccac	ggcaagcggc	119580
ccgtgtagt	ggcatctgct	catagccaca	gctaccagg	aggctgagcc	acaagactca	119640
ttgaagctgg	aggtcaaggc	cagcataggt	accataggta	gaccccatc	tcaaagttga	119700
acagtaaaaca	tatatatttac	tacaattaaa	aaaacaaggc	cgggtgtggt	ggcgcatacc	119760
tttaatccca	gcactcggga	ggcagaggca	ggcggatttc	tgagttcgag	gccagcctgg	119820
tctacaaagt	gagttccagg	acagccaggg	ctacacagag	aaacctgtc	tcgaaaaacc	119880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaccaa	aaacccaaca	catgtacct	ttctaacata	119940
aaatttttcat	tttttataaa	aattacagtt	ataattttta	gttgaaacata	atacaatgat	120000
aagtctccaa	cttgataatc	tgaggctggg	ggtgtagctc	atttagtaga	gtacctgcct	120060
agcagtcgct	aagtccgggt	agtcccttgc	ctgtatccca	gtgcttggg	aacaggcaga	120120
aagaggacca	gaagttcaag	gtgctcctcc	tcttcaggta	atgaggagtc	tgagccagcc	120180
tgggatgcgt	gagacacacc	agtaacagca	actactgcct	ggagtgtc	ctctggacca	120240
ggaacaagag	ttaagtgcgt	tactcccgt	actgtctgca	caatgatgga	gacagtacgt	120300
catctttaac	atctttggct	gagaagagaa	aacctggtat	tcctcagctc	gttctggcta	120360
agttcatgta	cttcacatc	atgcagttaa	tttgtaaa	accaatcctg	gcagcaataa	120420
ctctaattat	atagaacata	agatgtggta	attaggaaaa	gctactaatc	cacttaatag	120480
agtaaccttt	atctcttgca	aatctggtac	aagagacagt	cccaaataca	atgatggcaa	120540
gaattccaga	gcattgtaaa	tagcagcaat	tgcccttcaa	ttaacgcatt	gtaacgcagc	120600
agctgccac	aagacctcaa	atcaatcagt	ctatagctaa	ggaaaaatct	ttctaaagcc	120660
aaaaccattc	tacaaagcag	taacgtaggc	tccgtttata	ataacctgtt	ttggggccacc	120720
tgcaaatcaa	gctatcccag	gaagccagat	cgtaatctt	aggctctgct	ggctacacac	120780
tgggtcccaa	ccatgaggga	actagattac	agcaggtccc	gccctcggtg	acctgctcat	120840
agctatcatt	cttacagctc	attatggcaa	gtgagctctg	ggcagagaaa	aattcacaaa	120900
caaaccacc	aacttcccaa	gcaagcattt	tcttaacaaa	cacaaagaat	aaataaatag	120960
agcctgccat	ggtagtgcac	acctgtaatc	tcagcatgtg	ggaggcagag	gcaggcagat	121020
ctctgccagg	agttccatgc	cagcctggtc	tagacagttg	caagatcaac	aacgctatat	121080
ggtgaggccc	tgtctcaact	cccaaaccac	tgaacaaag	taagacgata	tggatcaatg	121140
caataatttct	cagttcctac	tgacagaaaa	tggacacaat	taggctgggt	ataattccaa	121200
tatgaataat	aagtatatata	tacagtacct	agtttaagtc	tgagcaagat	attatcgcca	121260
acagcacaga	tacaggacac	acacacagct	gcaagcgtga	aggatttaaa	ggcaccatc	121320

cctacagtat	atgcagcatt	gactgtctag	ttttatttcg	cacactttga	atcatccatc	121380
cattttttgct	caatatggca	gcagtaataa	aatgtatatt	tgcattttga	tgcatgggtg	121440
aattcttact	agcctgggct	gtgtttgctc	actaactcca	gtggacttct	gacgtaagag	121500
gcgctggact	agcttccaga	ggatgtaaat	ctaactttgg	ttctcggcct	cccctgaagc	121560
ctttgctgtg	gtgaaagggtg	ctgtttctga	agccacgcga	gtcccatggt	ggtttgagta	121620
acaatactcc	tggcttggtta	acaatgccaa	aaaataccaa	aaaaacccaa	accaaaccac	121680
aaaaaagcac	agagctcaca	tctgagccaa	aaaaaccgac	actccctatt	ttttgaagaa	121740
ctcacagaaa	tcaagaagaa	aaacaagcaa	acaaacagca	cataacaggc	taacaacaac	121800
aacaagtcca	cttcagaggc	caaaaaccaa	aaggcgaagg	ggctcttaaa	catttacaag	121860
gacagaccctc	actcagaaca	caagctatag	ccatgatact	gctcttcac	catcaattct	121920
taaaagacaca	gaagatgggc	tgtctggcatg	actgggaaga	gaccagtgtg	ctcactcatt	121980
gctgggggtac	gctgtaacaa	gcagaggaga	atgtctaaact	gtgtctgcca	atcacaggca	122040
cattcccagt	taaccagca	atcacttctg	gcaaaaaaag	cccactctgc	ccgcacactg	122100
gtgtaccag	gaggtaattc	actgcggctt	ctgggtggatc	ttttttcctt	cttgcaagtgc	122160
ttgggttcaa	attaggatca	agcacacttt	tatagtccaga	gcacggagac	aaaaggatct	122220
aaccacagag	ttgggttaaat	aaacaacaga	agaccacac	cagaactgct	ggggggcggg	122280
ggagggggg	ggaggggaag	gggagaagct	agaatcccag	cattcaggat	gcagagggtg	122340
gtaaacaggt	ttcaggctca	gtcgggggtat	aaggtaagac	tttattttcac	aaaaataaat	122400
aatttttttta	aagaggacag	aaaaaagaca	gcgtagagaa	ctagctatct	tttatgtaaa	122460
ccatgtgggg	tacaggagca	gcatttgcatt	ttgggttactt	ggcgaatagc	cctggtagga	122520
taaagaaacc	ttaaactgtg	ttatctataa	agggcagcag	tggtgtgcaca	atgcagtggg	122580
taggagccca	cccattgcac	gccatagttt	tgattacaaa	gtcacgcagc	tctactcaaa	122640
aattaaaaca	aacattaatg	cttggttaag	aaacagggtc	gcagagatga	tggtcagtgt	122700
gttaagagca	cagcgcccat	atggaggctc	tcagccatct	gcttctccaa	ttccaggggg	122760
agctaacgcc	ctcttttggc	cttcaagagc	actgcatgca	catggtacgc	ttacctacat	122820
gcccgcgaac	attcaaagaa	aaatacaaac	tgctataaaa	cccatatatg	accatcttaa	122880
gattctttca	tttttttgag	acagggtttc	tctgtgtagc	cctggctggt	ccagaacttg	122940
ctctgcagac	caggctggcc	tccaaccag	agatctgcct	gcctctgcct	ccacagtgtc	123000
ggaattaaag	gtatttaaca	cacacattat	acatatcttc	cttcttcttc	ttcttttttt	123060
aaagcgtttt	gttggttttt	gttaagtgtt	aattaaaaaa	ctacatagtt	ttataggcaa	123120
acataattaa	aaatgccaat	gtgaaataaa	taatataatac	atatataaca	ttctgtaata	123180
gattcactca	cacaacttat	atacttaaat	acaattttca	caataatgaa	aagctttgaa	123240
atgaagactt	ctggatacat	tagaaacgta	ccctgaaaat	cgcaaataac	ggttttcatt	123300
tctttgtgtc	agacattagt	gtgagtgtct	aaacttgcat	aaaggctctc	ttctctatca	123360
cttctacact	attgcagtgg	ttctcgctaa	acctggaacc	tggtggtcct	gttcccttggc	123420
tggaactaca	ggcagcaagt	cccagcaatc	ctcctgtctc	acctttcttg	gaaccaatgt	123480
tataagtgtg	tgtgggcact	agccttggtt	catggctgct	gggactagaa	ctctggtctt	123540
caatattagg	catcaagagc	tcttaactgc	taagccatct	ttctaccctg	attagaattt	123600
cttgaagcaa	aagaaactca	cagatggtca	gagtttacac	acacacacac	acacacacac	123660
acatacacac	acacacacac	tcacacacca	aggcttagtg	accactgtga	aaaggggaagt	123720
gcggtgagga	actgtaaaaa	taagggtgtca	ggaaagctct	gcacaaaatg	gtgtcctctg	123780
gacaagccag	ggcctctgcc	ctcatcagct	cttagtaact	atggttgctc	acagcaaac	123840
atgccaggga	ccactctaac	atggagctgg	gatgggctcg	agaggccctg	ttattaaagg	123900
agaagctgtg	gagagtgtat	ttgatggatt	ctaaagtagg	gggaatcgat	ttcctttatc	123960
gatattgggtc	agccatgctc	tagtgagtgg	ccccacaccc	acctatgagt	atgtgtggac	124020
agcacacacc	ggacctggca	agtcaataaa	acaaaaacaa	aaacaaacaa	ataaacgttc	124080
tggtccacca	tagtggctgc	tctgtggttg	gagctgtccc	gcacttatgg	caggaagaca	124140
gtggctgcaa	agaggacaaa	agtctctgga	actgatcaac	tctagagtct	gcttggttatg	124200
agaactggga	agtaccgct	gggacagaa	cagactctga	aggtgatcag	gacagagatc	124260
acagaaggag	agactggtta	tctggaagaa	tctgaaacat	aactcgacgc	atactgggtc	124320
aaactgggtc	ccatcactac	aacagcagta	attgaattgg	gcacaacatt	cagaaaacag	124380
aaaaagacta	cagagtacgc	acctgggcta	tcatcaaccc	aggcgattct	ggcactattg	124440
gaagcaagcc	agactggaga	aaaggaaaca	aaaagtattt	taacaaaact	tcccagagca	124500
tgttaaaaaa	aaaaaaaaga	aagccagaca	tgggggtgca	cgtttttaac	cctagcactc	124560
aggaggcaga	ggcaggtgag	gcaggaggat	ccatgagtct	gaggccagcc	tggtctgtac	124620
agtgggttcc	aggaaagcca	ggcaacaaag	aaacctgtc	tcaaaaatca	ctgactgggg	124680
agagaggga	tggatccaag	agcaagagag	agagagcaga	gagagtgggg	gtttggatgt	124740
cagcgttatt	aaatgacagc	agaaaagatg	ggcccagaca	atgacatccc	agaaatggca	124800
aagatgaaaa	aataaacaca	agttctaaat	atcattttta	taatggctgt	gtgtctgctg	124860
gctcattctt	gttatccaaa	caaactaaag	caggggtggc	cgggatttac	ggccagccag	124920
gactagagtg	agacactgct	ttaaaaaagc	aatagatgca	cgctaaccat	taatcagcgt	124980
aactggagtt	tgagggaggg	agtgggccc	ggaagcctgt	gcctataaac	ccaacctgcg	125040
aggcctgaag	cccgaaggtt	gaactgagaa	catctcaaga	caaagcacag	gcacaatctc	125100
ttacaaacag	tttaaacaca	ataccagcaa	taaattgtca	gctttatgac	agataggctg	125160

acaggcatac	cacaaagatc	gggaagaaga	aacgggattc	ttgcacaaca	ttttacaaat	125220
cgacacagct	gagcctagt	acaggccgtg	ataagctcaa	aacatgcaca	ctgcaaacac	125280
cacagcatgg	ccagagtgc	agggtcacag	caatacacga	acagagacag	taagaaacat	125340
tggaaaaggg	aagaaggcaa	agagcaccac	gagagaaaag	ttaacccgcg	attccatatg	125400
aggggccagc	atccctctcc	cacctgacag	agaaaaccag	gaggacgcag	ctgaaacact	125460
gtacaactat	aaatactcct	tcctagtgtg	gacagagttt	accaaagggt	atcgtaatct	125520
gaagcacaaa	cataacactg	taagttacca	gaagtttaga	cacgttgaga	atttaataga	125580
agttgagaca	tcaaacacac	gtttgatcct	aggggaatta	aattattggt	aacagaaaag	125640
ctctttacaa	agtctccaca	tttataaacc	aaatacactt	gtatcagaat	tgttgtctga	125700
gccagaggcg	gcggcacaa	ggtggctgag	gcagcaggat	caagtctgga	gaagcccctg	125760
ggtccaggga	gcactagtga	attcaagtaa	acattaaaca	ttaaataaga	aagacgggatt	125820
ctacaggagc	ctagacaccc	tgtaagggtc	gtactactaa	gataccaaaa	ccaaactgac	125880
tcagggggct	gatgagatgg	ctcagtggct	aagggcatgc	actgctcttg	cagaggacct	125940
gagttcagtt	cccagaaccc	gtatcaggca	gttcgcccac	ctgtaatgta	ccagagaccc	126000
taaacatttc	tatcctccaa	gcacagacac	acataaacat	aattaaaaat	aatctttaa	126060
aaaaaatctc	tcgtggacag	acagcaaaaa	atactgaacg	aaatctaata	aggtagactc	126120
agcaaaagca	atccaaatat	gcgaaaagga	taatacaatg	caaataccaga	cttatcccag	126180
aaaccccagg	tcgcttttag	atccaaaaaa	ttcaatcatc	ataattcacc	gtatttagcta	126240
accgaaacag	aaaaggcatt	tgattattaa	taaatacagg	gaaagcattt	gacaacattg	126300
accattcact	cttaataaaa	atgttaccaa	acaggagaaa	gaagtaaaat	ctcctcaacc	126360
cgatgaacag	gaacacaaac	tggctgggca	cagtggcaca	tcctgaaacc	ccagcaccca	126420
gatagctgcc	agactgggct	ggcatgagaa	ccaagggtaa	gaggcaccca	aacacttaat	126480
gatccaaggt	gctgcctccc	ccccccccc	gtgtcactta	aatccgtaat	taaatctctg	126540
ggaggagggt	tttgacacag	ggtctcactc	tgtggcacag	gctagcctga	atggagtcaa	126600
tgctggcctc	aaaactcttc	tgctcagtt	ccagagttag	gggaaacaat	catgagccac	126660
ctcaccaggt	tttaatgcca	tttaacagta	caagggttaa	cgctgcctga	aaagaacaag	126720
acaaagaaag	atcacacaga	caaacatcac	acagggtgct	tttgagacta	ggctcttagt	126780
gatgtaggct	ggtctcaaac	ttgcacaaga	aaagaatact	tttgccatca	agatcaagggt	126840
tgctatcacc	catggtggca	tttcggaagc	agaggcagga	agatcagtag	tacaagggtca	126900
tcctcagctg	ccatgggtgt	gaggccagcc	aaggcagcac	gtgagccagc	tgtggaagca	126960
cacatctgta	accccaccac	tcaggaggct	gaagcaggaa	gttcaagcca	cacaagagcc	127020
acttaaaaaat	aaataaaaaa	ataaataagg	ttggagagag	ggnnnnnnnn	nnnnnnnnnn	127080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127200
tgctcatctg	aagggttgaa	atcactgagg	aatggctaata	gacaaaactg	cttaaggaga	127260
attttattaa	aaaacaaaac	tccttctgga	ggagtactgt	gctcccttca	gcatttaacta	127320
ctgtgccagg	gaaccactg	ccagcagaca	cccaacttca	aggataatca	agcttgctac	127380
gtaacatggc	tgcttaacga	gacaccacac	acacttctct	tatgctgggc	atcacctcta	127440
gcttacttat	aatatcaaac	ataagtagca	tgtgagcgct	acagggcctg	gccaatgaac	127500
agcgatggga	aaaataaacct	gcacatgggt	agtaatgggc	aatccactcc	ccagacactt	127560
ctaacctcag	ctgcagactt	gtgggtactg	agaacggacg	tgactaagtc	aattcacaga	127620
agcagagccc	tgtgcctgtg	ccagggaagc	gaggacgctg	ctcaatgggc	agagcttcag	127680
ggcggaagga	tgtgaaagt	ggagctaggg	ccgatgacg	tgaatatggc	taatgccata	127740
gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	gtatgtgtat	gtggtgtgta	tgtatgtgtg	127800
tgcatgtgtg	tgtgtgtgtg	tgtatgtatg	tgtgtgtatc	caaaatgggt	ataaaaaatcc	127860
catacaatgg	ctatgcttat	gtgtctttta	ccacaagtaa	aaaattttta	gtaatcttag	127920
gaacacattt	ctaaaatttg	gaatacttgt	tggggaagct	atctgagccc	cagcatgaat	127980
caggaaatgg	gtaggagagg	caggagagat	ggcttagtg	ttaacacaca	catggtggcc	128040
ctttcctggc	accaactagg	tcactcacaa	ccagatctgt	acagcctctc	ctggcctcct	128100
caggcaccag	gcaggaaagc	ggccccacat	ccctgcagg	aaaacatttg	tactgaaact	128160
aaataaaaaa	ggagcagaag	ctgggcctgg	aagtccttta	cctccagcac	cgggaagcag	128220
gtggatattt	tgtgagttcc	gggtctacat	agtaaaaact	tgtcgccaag	taaaacaaaa	128280
caaaaactgg	gggctgggtg	gatggctcag	tgggtaagag	caccgcactg	ctcttccaaa	128340
ggtccaaagt	tcaaatccca	gcaaccacat	ggtggctcac	aaccatccgc	aacaagatcc	128400
tcttctggag	tgtctgaaga	cagcaacagt	gtactttacat	atattaataa	ataaatcttt	128460
aaaaaaaata	cctttaaaaa	aacaaaaaaa	aaactcccac	taaaataatt	ggaaaagtcc	128520
agggaagcc	acactcgatg	gcgcaagcgt	gtgatcacag	tattctgtag	gtaaggcaag	128580
aggatcacag	cgggacggag	gccagcctca	gctacacagg	ccggtctggg	ctacagtgtg	128640
agaccccggt	ctcaaaaaca	aacaaaaaaa	aaaaagcggt	cacattatca	tatactcaag	128700
gccacaaaac	acottctctt	tcccaatcct	tgaatgctat	cgaatttggt	tacttttttt	128760
tggtaatatt	ttgtttgctt	ttcaaggcag	ggtttctctg	tgtagccctg	gctatccctg	128820
aactcactct	gtagaccagg	ctggcctcta	actcacagag	atgtaattct	ttttcaaagc	128880
tggattttga	tttgggggtg	tgtgggtgat	accacaggga	cacttggggc	ccaaacaaaa	128940
ccaaagaaaa	aacaccccc	ccccaaagta	tgaacatcaa	ctgtatataa	aactaacagt	129000

tcatataagc	taagtagcct	gcaggtacat	ggttatggaa	acagctttat	catacagact	129060
tctttcagta	gatgccatth	gagaaaaaaa	aaaaaacaaa	acattctttg	gattttcaac	129120
atgtaaacga	aaatatgtta	aatcttataaa	aaccttaaaag	cagacgccac	tgctcttttg	129180
ggtcagggaa	gcacggctgc	aggcccccag	agcagcacat	tccctgaggg	aagccttctg	129240
cgctctcacc	agcaccaaga	acagccaact	ataataagct	cataaaaaat	cctgaaaggc	129300
tgcccaagcc	tgaaaatctg	atatgaaaac	agagggcaag	acagataaaa	ggcaaaactat	129360
actttataag	tcaattccag	ctatctgtgt	ggaaggactt	ctcgggacgg	actatctgct	129420
aagaccttgt	gggtgattta	aacagcgtgg	agggcagaaa	aagaaaggaa	agtgcaggaa	129480
acaatggcaa	aagctcctcc	tccgtggcct	ctacacgcat	ccaagcgata	tggagggagg	129540
aggaaggagc	actatacaca	ggtaccaatc	cccacgaact	agaaaaacaca	gcttttacta	129600
actagtctat	tttttttttt	accttagcta	gtgtctttct	tatgtttggc	ataattttctg	129660
ctctgcatta	aataaaccta	gtgtataaga	aggcaaatga	gtaaaggtaa	agtcagttaa	129720
tgttcatttg	ttttgactgt	gtggtgtgca	tgccaggcgc	attcgtggag	gtcagaggac	129780
aaccgcatgg	agcaggctct	ctccttccac	caccccaggg	gtcgcagggg	tcccagggat	129840
ggaactcggg	tggtcagatt	tgcttggaac	gggcttcaat	cactgaggcc	ccccagggt	129900
cccaattcac	ctgtttactc	taatgtatat	atttttgaga	cagggctctg	ctatgtagtc	129960
caggtagctt	ttgaactcac	tgtggacact	agctggcctg	gaggtctgag	tgaacctcct	130020
gtttccaccc	aagtgtctgg	acaacagaca	gcaaccatca	aggcaaaatg	aagtcttccc	130080
ttcacaccaa	agtggcttcc	tatgaccttc	tctgacaac	cccaaacagc	aagtgcctcg	130140
atatgcactt	gcgtttctct	ttttttactt	caagctttga	ctccactttc	ctaaaagggtg	130200
tttactaggc	tgaagcacac	ttcaggagac	accctgtagc	tctggagtga	ccggaaacac	130260
accagctctg	atgcaaaaac	aaaaaacagc	gatggggatg	gggcaaaggt	ggcttgtgct	130320
tctgtgcaat	ggcattccct	gacgcattac	ctcacaccag	atacatacag	aaaagacaaa	130380
ggtaagctct	cactcagtgc	gcgcattccac	cacacccaaa	gcaagcattc	aacagctcag	130440
tggtcataag	ccagcattat	cacagcgcgg	acaaacacaa	agccaacctt	ctctttgagt	130500
tcccaaggt	acatagtgtt	ctgtccaagg	atggcctctg	cagactcctg	gaaacacgtc	130560
acccactgg	tctcttgaaa	atctgcaata	tttgccctaa	aaacacaact	ctattatttc	130620
agcgaataaa	taacaaaagg	aagacacccc	aaggacgagg	gaagattatc	ttttcctact	130680
taaagtaact	tccaagggag	aaatgacatt	gccgccacac	acagcccccg	actcttttgt	130740
taacggttcc	cccttatgag	gaagtattca	cttctgtgtc	ccccgtgttt	cacttttagt	130800
acagtattcta	atgaattaca	ggagccactc	aataccttta	tcataaagtg	ggctttgtgt	130860
taggtgatcc	tgaccaagcg	taggctatgc	aaatgctctg	agcacactta	ggggaggctg	130920
ggccgtgcta	gccctgctac	atgcaggtaa	tgctacgcca	gtcttacgcc	acaggaagaa	130980
gcttaaaaatg	tggtcgaact	caatctgcta	gaggtgtgag	tttagacctc	gggagacggg	131040
ccaagtaaac	ttccgcagat	gtcggagcat	cagacagagc	tgtccccctc	tgatgcaaaa	131100
ggcttgcgac	ggcaagtttt	taatgagcct	cagtggtata	agtgggttcc	ctctctcctc	131160
cctcccttat	caatatgctt	gacatctttt	agttttttga	gacagcaaat	aacgtagccc	131220
aatctgcccc	catacttact	gggcagccac	agctggccct	gaactcctcc	tgcttctcct	131280
ttaccccgta	acaagtgtct	gggttacaga	catatgccgc	catgttcagt	gtaagtgacc	131340
gctgcccagc	agggcaaaag	cttccttatt	tacaaaagcag	cagccaagcc	ctgcagccca	131400
ggcctatctg	atttctctcag	cacaccccca	agggctctcac	cgatggcagt	cagtccatga	131460
acaccgtagg	cttctcacaa	tccacactac	tcaccgataa	gatcatgcgg	tatttgaaat	131520
tgggaaattc	ccggtcacac	ttctcacagc	ggtacaaacc	attctgctgg	tcaatcactt	131580
tcttattgca	gtcctgggtt	gggcaggcct	ggtacataca	gttctctttg	cggagaaaca	131640
ccaccgtctg	cacagtgtct	aaatagtccg	cctgcaggcg	aaaggaagac	cgccatcagc	131700
aagcaacaca	ggtctggaac	aggcaactca	aagctgctct	tccttgcaag	tggtgagcgc	131760
gtgtacatcc	tagctcccac	gctcacgagc	gatatgcaga	atcactaact	ctgggtttcag	131820
aaatcacacg	tgctacacgc	agaaccccaa	gaagtaacaa	accggcactc	accgacctca	131880
ctttttccat	tttgagagcg	goggctcact	agctagcctt	gaactcagag	tttgggtctg	131940
ctctgtctct	atagagtgc	tggtaccac	acctgggcac	tttgttttgc	agacagggtt	132000
tctccggaac	tactctgtga	gaccaggctg	gcctcaaaact	cagaaatcca	cctgcctctg	132060
cctcccaagc	gctaggatta	aaggcatgtg	ccatcagcgc	ctgacccaat	tctttttatt	132120
tatagtattt	atttggtaca	tatgtgtatc	agtgtgtatg	acgtccatat	gtatatgact	132180
gtgtgcagtg	tgccatgtgt	tgccaggtcag	aggcaactc	tcaggagtca	gttctctcct	132240
cctactgttg	cgttttggga	actcaggttc	caagatagca	ggaaaagtgc	ctttaacagc	132300
tgagttatct	cgacactgac	atgtaatgat	tcagtgtgca	cacagtgggt	ttgcatgtag	132360
tcatataaac	cacagcgtgc	atgtgcaggc	taggaacaaa	cctgtgggag	ttggtgatct	132420
cccctcacca	tgtgagttag	ggagagggaac	tcaggctgtc	aggctcgggtg	gcagtgcctt	132480
tactcactga	gttcccttgc	tggccaagca	ttattttataa	gatggtgatg	tctaccctta	132540
tcttttagga	tcagtaagtt	tttccccaac	acaagccaga	aaaacctttt	aggcccaatg	132600
agccattcag	tcagttctcta	ctgcccactcc	tcaacctgtc	tgtggggcag	gacaagccac	132660
agacaacctg	aagacggaag	gtgagccaat	aaaactttac	tgacagtaac	agccagccag	132720
ctcacaggct	tgacaggcaa	ttcttggaact	gaatgcgttt	aagagaatgc	agaattaccc	132780
atgactaaga	tcttctaaat	ggaaaaatgt	ctggttaagt	aaccagcaaa	ggagctaagt	132840

cacgcaagcg	gtggatacct	gctttgcctc	tgaaccctgc	acaggttttg	gtttatgttc	132900
aatcatgtca	agtacctaca	aaatccagat	ctagcctgaa	ttcaaagata	ggctacctac	132960
cctgcecccg	gacccccacc	cggggtctca	ctgtgtagtc	ctggctgtcg	tagcgtcttc	133020
tatacagacc	agctgttatc	aaattcagag	acctcattgc	ctcccaagtg	ttgggattaa	133080
aggcatttgc	cactatgcct	ggctctcact	ggttgtacca	gaggagcaaa	acaatgtggc	133140
catttaaaga	gacgaagcta	gaaaccagtt	tagacagctt	tggggctgtg	ggtgtagctc	133200
agtggaagag	cttgcttagc	atgcacaagc	tgttggtttt	aaccctcagc	gtgacagaac	133260
cgaatacata	agaaggctta	gaggaggggg	tgggtggtga	gagatgggtt	agaggttaag	133320
agctgggtgc	ttaatttccc	agtccccaca	tgggtggctca	caacatccat	aactgcagtt	133380
ccaggggatc	tgatgtcctc	ttctgacctc	cttggggaca	tgcgactcat	ttggctcaca	133440
tgcaggatgc	ctcgggccac	tatgctggct	cagggttgaa	ggtgcttgct	gccaaagtgg	133500
gtggatttga	gttttggtccc	tgggaccac	aagggaagag	ttgacttcta	tacactgagg	133560
tggaaatgc	atgctctacc	cgcaaattaa	aaacttaaaa	tttaaagagg	aagctgtaga	133620
gaaatagctt	ttaggaggat	gcctaaggaa	cttctctgcg	ttttcagggtg	agattcagac	133680
tcaaagccca	atttaaaagt	ttgagtgtctg	tcacgtgttc	tgtatgcca	gttctggctg	133740
tccgttgtct	gtcttttagtt	taagagagca	actgggtgag	aagtaactga	gagtctagcc	133800
gatgtttaat	tctcaagatg	tctgtgata	gcattataag	ttgctgtgga	tgacagtgat	133860
ggatcgagca	cacagtgaag	tgagacagtg	aggaagaaaa	cctatatattg	atacgaggac	133920
aataatggag	cagtggcgag	cgatatatttg	ggaatacaat	agaaataaat	gattcaataa	133980
tccaagagaa	cgcaggtttag	accaaggtag	gaagaggatc	actgggctgg	agagatggct	134040
cagtgtgttaa	gagcactgac	tgtctctctg	aaggtcatga	gttcaaatacc	cagcaaccac	134100
gtggtggctc	acaaccatcc	gtaataatat	ctgatgccct	cttctggagt	gtctgaagac	134160
ggctacagtg	tacttacata	taataaataa	ataaatcttt	aattaaaaaa	aaaaaaaaaa	134220
aggaagagga	cacttgagca	cacattgaaa	tgaagatgc	aactgaaatg	caataccaga	134280
gccagtctgt	ggcatggcag	caggaattac	agtctgttca	aaaccagca	cggcctgtag	134340
gctacattat	gtccctctac	actggccgtg	actgaaaagc	agcaacgtgg	taccctggag	134400
caggcctgag	gtcccaggta	aggacacagc	cctgacctgg	cactaggaaa	cagggtgtaa	134460
aactcaagcc	cctgccgttc	tctgggtgtc	tggagcgagg	ggtgcgaggt	accttgtctc	134520
cctggcccgag	gttctcagat	ttagcctcat	gcaaagtgtt	ccagttgggtg	ttgccccctc	134580
cggccccctc	actcctgtgg	tcagagatgg	aaacaccatc	taaggcttgt	ccttctgagt	134640
caaacctagg	ggagaaagaa	acaaacgtac	tgtcgacagg	aacttcacat	ccttctcaaa	134700
tgtgctgtga	atgcaccagc	cccagactgc	gccccctcgg	ctctcaccta	ggtctcagca	134760
cacctagtcc	actcaagaaa	tgcaatgccg	ttgtctcttc	catctcgtgt	ccttcaatgc	134820
ccctcctcct	ccaccgcgaa	ggtcagacaa	actgccaaag	taccacagac	cctttcccta	134880
aactgggctc	acattgacag	cagccatgaa	caaaggcagc	aacagcaagc	cacactggca	134940
cccctcccc	caccctcggc	ctcaagcctt	caccccaatc	accaaaatga	accagatgaa	135000
aagggacatt	tttgtttcat	gagtctgaga	aaaatgtttt	catgggagaa	aatcttttaa	135060
gcccaatcag	ttatagaact	caatcagaag	gcattatctg	ctgtgttaga	gatttgcatt	135120
ctcagtgaga	atttgtttat	ggaaatatca	aggttaaatt	cattttatatg	aaattattat	135180
taaaattgac	tatacttctt	gaccatctat	tggctaattc	gaaggaaaag	aggccaggga	135240
gaaggtagca	aggacaggct	agtggcagag	gacagtgagc	ctgagatgaa	aacacctcaa	135300
ctacacgatt	ggttgtaact	aaagctgccc	ccacacagga	ctgactccct	aaggagactc	135360
ccaacagagt	agcgggtggc	caaggcaagc	agtcacatcc	gtcagaggac	aaggtccata	135420
aatccgtctc	accactacta	gccgctcaag	tttaacctca	cacaggcatc	ggcaaccata	135480
gacggctcat	caggagtga	cagagccagg	agacaggcag	cctctcttat	agtaaaaggt	135540
ccactcttag	agccaagatc	taaagatcgc	taatccttgc	ctgggtaggg	ggggtatact	135600
gaacttgtag	atcaatgaca	gtattttgtt	cgggggacag	acatgcctgt	ttgtgatttc	135660
tgtgatcaca	gttactgaca	gtgaaatatg	aacaccttaa	tgtagacaca	caattaacct	135720
ggggttgtgc	tcagagtctg	ggagaaaaac	caaccaacaa	accagaccaa	tgccatggct	135780
gtctctgggtc	agacttctgt	cctacaatgc	aacgctagct	atcagcacag	ttacacagcc	135840
agaccacac	cactgaggat	gccttaagtg	acagacaccc	cagatggatg	ctctaagtag	135900
taaatattgt	tattggttta	catcagactc	aaaacaaaaa	tgaggagcgt	gaaatatgag	135960
ctgtttttct	tgagtctctt	aaattctacg	acacagctgg	aaaccaaca	tgcccaccct	136020
ggtactacag	ctattcaatt	tccttaagct	tgggctagg	aaagtatttt	ctttgaccac	136080
ctcaatcctt	ctacacaaca	ccctcaggat	gaagtgtctt	tccaaggcag	agtttaatgc	136140
cttaaaaagc	tagatatgg	gtggagggtc	taacaagctc	tcctcatctc	ctgagtgcct	136200
tgggaatgac	atgttacagg	caacaagag	tccattgttc	tcactgtact	ttaccgagtc	136260
cagtgacaac	acaacagaag	atgtgaccag	ctccagagac	tcagcagagt	aagaaagtac	136320
attctacccc	cacatcagct	gcccacagtg	caaactgaag	atgctctcag	cactgttctt	136380
cagcgcgcag	ctgttaaagc	cttgccgtac	ctggcgggat	caaggctcagc	acttagattc	136440
taattgcttt	ctctggctct	ctatgttcat	caggggatcg	tgggtttgtt	ggtttaagag	136500
ggctttcagg	ccaccatctc	agaagacata	atgcctcgaa	gaagtatagg	aaccactcca	136560
ctcaatagca	aaggcttggc	aaaacagctt	ggcagcagtc	cacatgctga	cagtgtccag	136620
ctctggttta	ccagagcccg	agcacagata	accctgagc	tgagttgcac	aaaacctacc	136680

agccacgaag	cttataggcc	tctgggatgt	caggattcac	aatgacagt	ctggatgaga	136740
ggaccgagag	gctccgtcca	ccgaagtcag	agactcgggc	tcctttgatg	gccatcacgg	136800
gctgccgaga	gccgtcaaac	ttgtcagcct	gcaaatcaaa	cgacgcgaca	catcactgac	136860
aagaagaatg	tacaggatgg	ttttatggag	aagagtcagt	caatgtctgt	ggactctaag	136920
acagacctcc	cactcgggaa	ggagcattgc	actacaagaa	gctgcaataa	ccgatcatct	136980
cacacagcga	aggtcttcaa	atacttactg	gatagcacia	ctccctgagc	taaagcccca	137040
ccctcaggac	tcgggtcagt	gcaaggactc	acatcttctc	cccacagagt	tgtgggtcacc	137100
accttccctg	acatgtccat	caaatagata	tttctcttag	caacttctct	gttgttcgac	137160
ttcactgtga	ttttaatcga	atcttcatag	ctcttgcaga	ttccaatgat	gtctgaaaca	137220
agaacactt	tgtaagagct	cccatcaagg	ctcctcttta	aaacacggca	gggacgaaag	137280
gcaagcacgc	tagtcacgca	tcaccccaaa	cactggagac	aaaaatcacc	actctttgcc	137340
ctcaaccctg	gacgcacctc	ctagtgcgtc	tttagccttg	ctctctaggt	caccgatccc	137400
tgtgaaatca	aactgaactg	tgggtaagt	atggccatct	tcacagggaa	ggacagaagt	137460
ctcattattg	aaggtcatct	catagtcatt	tttaacagcg	gagaactgtt	tgttagcgat	137520
cttcagggcg	ccctttgaga	agtaatacac	ctgcaaaaga	ggccaagtca	gggcaagcta	137580
ctcagtcctc	caagcccgcg	cccacaatgc	agccttccaa	actgtgctca	tggctctcac	137640
agttcagcct	gtcagctctg	actcaacacc	gtcttcttac	cttgttcact	tcaataaggg	137700
gaaagaactt	gtccacttgc	tcattgaaag	cagtagctct	gatttcaccc	tgcagaagca	137760
aggagagca	cattagaact	gcgctgctag	gccctcgctc	tcaacagcga	gagcaggcca	137820
ttttgacagt	tagacaccat	cctcggagca	acagtccaag	tgtaaagaga	ccctcacagg	137880
gctgcagaca	ttgcagaatc	ccacatatac	agttacattt	tcaagctatc	agctctggta	137940
aaacaaaagc	agctcagcca	cccagcgcta	cagttgtaag	ttaacacata	aaacgaggga	138000
gcctcgaagt	gactctggaag	acaaagggtg	ttcctgtgcc	aacctggcca	ccctagttta	138060
atcccggaa	cctcgggatg	gagaagacaa	tcgccccacc	aagttattct	ccagcctttg	138120
cacacgtgtg	catgcataca	cacgaaaata	attcttacat	cttatcaaaa	cattttttaga	138180
aggaatagct	taatattccg	ataatgaaac	ctaattattct	gggtcccag	tgatgggctg	138240
gaagcccaca	ggaaggcggt	ctgaactcca	atgattcagc	actcttcctc	acgcccgcga	138300
cctaagtgc	ggcttacttc	acagtcagga	gctggtgcca	aaaatgttcc	aacaagtcat	138360
gctttcaagc	tgactgcact	gttacttttc	tgtagcgcaa	tatgtttacag	atggataaga	138420
aacttaacta	aatggctaaa	accttaccct	gacaagctgc	aggaaattag	tttctgtatc	138480
tgaatctgcc	atcacagtta	agatatctgc	tgtgctacac	cgtgacaaga	ggaagaagag	138540
ggggagagag	ccccaaagct	ttgcccttcc	cggtatcatg	gcttatttta	cgtgtgtggg	138600
tgttttacca	gcatgtatgt	gtgagcatca	ctgagtgtct	gatgcctgag	gtgaccagga	138660
gaggtactgg	atcactgaca	cgggagtcac	agatggttat	gtgctactct	gttgggtgctg	138720
agaattgagc	ccaagtccct	tcaaagaaca	gcgagtgcct	ttaaatgctg	agcagctctc	138780
ctggcccctg	cattttgctt	taagttaaagc	ctagttagta	taaatgatca	gaagccctcc	138840
ccgcagacta	gttaaagctg	agtagctgct	gtccttctg	ctggaggcaa	accgcctctg	138900
ctcggcaggt	atcaccccag	ggcttcaact	gtgcccctag	caatgtaatt	agagcgctgc	138960
agtctctgca	gacggagact	atttacagtc	caccaatctg	tattgtctaca	gacgcgtccc	139020
tttaggagca	ctttatctca	tgtctgaccc	tgtgccccag	atagcatcaa	aggtctccaa	139080
cagaaggaaa	gccacaatga	ccgccatggt	cctcggagct	gagccccacc	cgnnnnnnnn	139140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139260
atctctatat	atgaatgagt	atcctgtagt	tatcttcaga	cacaccagaa	gagggcatca	139320
aatcccat	acaggtggtg	gtgagccatc	atgtaattgc	tgggaattaa	actcaggacc	139380
tctagaagag	caatcagctg	tcttaacccc	tgagccccct	cccttttttt	tttttttttt	139440
ggcagcttat	ttttaattgt	tttaaatcac	gtatatgttt	ctgtatgtag	ctatgttcac	139500
ataagcgcag	gcactcatag	aggtcagaga	tgtaagatgc	ctttggagct	ggacttaaac	139560
atggttggtg	cccatctgat	atgagcacca	agtggaactg	ggctctctgg	accgccttgc	139620
tgccctccac	ctagaattac	agtatatgca	attagctgct	cagccacccc	ctcagtcctc	139680
cttagtactt	tcaacatagt	gtaaatcgga	gagcactcct	tggtgcaaag	attctgtgtg	139740
ttctatcatt	ctctacaccc	aggcttcatt	gggacgtgtg	atgccagcga	tgattcattg	139800
gcgtgaagta	tgagttagag	tcagacaatc	agtgtctctc	gctctgagtg	ggtctccctg	139860
ccctctgtgc	tctgagacag	gggaggattc	tcctgttgga	tcacatagtg	catagagccc	139920
tgtgtgcaca	gcgctctcaa	ccattttgtg	ccattttcaa	tacagtgact	tcagagtctt	139980
ccttgaaata	acagatttct	ccattttgtt	tgttcccttc	tttcgtcaca	aagacctgag	140040
gatgagatgt	ttttgaagga	actttctagt	agttactcgg	tggaaaagga	caatgatgct	140100
cccctcttct	acagagaaga	aggaaacagg	aaattccaag	aaaaggagta	cacagatgct	140160
gcagtgtgtg	actctaaggt	aacgtgcgtc	caaccagcag	ttgaacaagg	cggagatag	140220
aggcaaagt	cagactcata	accagtcatt	tgttttgctt	tggttagttg	ggttggttgg	140280
ttgtttgtt	tggtttgttc	ttcttcttta	tttagtcgga	tggatttggg	aaaaatgtag	140340
tagaattcta	tgtgtgatta	tattcacaaa	caaaagactg	tttatatgtc	ctgaattttg	140400
gcaagattct	tcccatattc	cttcagaccc	atagcagcag	caagcttagt	ggccctttgc	140460
caggtattct	cactgagctg	tacagcattt	gtctgaagaa	ctgatgaaca	tttagcttta	140520

tgctcctatg	ttagtaaata	gctaggtttt	tagctagcca	ccttgctagc	ttacaactag	140580
aatgcctctg	gctgagtttg	gtggcacttg	ccttttagttc	cagcactcag	gaggcagaag	140640
caggtgaatc	tctgagttca	aggcagcctg	gtctacataa	caagttccag	cctggccagg	140700
gatatgtagt	aagactctgt	ctcaataaag	taaacagcca	gaatgactta	aattattgcta	140760
aaaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	140820
agaaagaaga	aaatgagtgg	atttgccata	gcctagtcaa	actagatttt	cttggttata	140880
tgtaaacata	ctactattaa	caacaacaaa	tactttacac	taaccatgac	aagctatatt	140940
tttaaattat	tattttatat	gtatggtgtg	ggaagctgtt	gcagagtggc	agttggctac	141000
tgctggccac	cacacatata	taggcagtga	aggttctttt	gccaaagaca	gttaaccaat	141060
catagtgtgag	acacgcctct	cctaggccta	tgtaagcagc	accagttctg	ggctcagggt	141120
ctcttcgcct	ctacaatcaa	gctctcccaa	taaacgtgtg	cagaaggatc	ctgttgacgc	141180
gtcgttcttc	ctggccagtt	gagcgcgcac	aagagtatgg	gttttcacct	acattttatgt	141240
gaactgcatg	tgcactttgt	acccagggag	cccagaagag	ggcatcatat	cccctggaac	141300
tagagtcaaca	ggttatgtat	gggttctagg	aatcaaacc	atgtcctctg	gaagaacagc	141360
cagtgttctg	gatttaactg	ctgagtcata	tctccagccc	caccatgatg	gatttaactg	141420
ctgagtcata	tctccagccc	caccaagctg	gatttaactg	ctgagtcata	tctccagccc	141480
caccacgctg	gatttaactg	ctgagtcata	tctccagccc	caccacgctg	gatttaactg	141540
ctgagtcata	tctccagccc	caccacgctg	gatttggaag	cagaactgag	gctttgaaca	141600
ctgtgttatt	ctaactcttg	cttccttgga	accctgagaa	aattccttct	attggctttt	141660
cagagatctg	tatgggctta	aaacaagaat	atgtccaact	cttgatctct	gattttatat	141720
attaaaagaa	tagagtgagc	ctaggagttt	ggcacactcc	tatgttccaa	gcttcaggaa	141780
gggaagaagg	tcaggagtgt	aaggctagtc	agtcttagct	ttgtgacaag	tttgaggcta	141840
acgtgagctc	tatgagacct	tgtctaggag	aggagaaga	aaggggtaaa	gtgaggtgtg	141900
gctccacgct	accacttgcc	taacacgcata	gacgcccaag	gtttccgttc	ccagcctggg	141960
ggttcagagt	attgtgcaca	ccctgtatct	ttaaaagtca	atagtcaacc	cttgggaatta	142020
actttccgaa	aaatattaaa	gctacatcat	ccactctaca	aacttgtaaa	agcccttctt	142080
tgtaatgggtg	taagaaagta	tttggttcca	gttttatggc	ctttgcatcc	taaatgttac	142140
cccatggaaa	gttgcttaat	gaaaacaggt	aaaaataaga	caggagggac	tgcaaagatg	142200
ctcaggctca	cgagtgttta	ctgcccttgc	agagaccatg	gcaggcagtg	gggtctcccta	142260
ctgcagggaa	tccaaaaccc	tcttttgccc	tctgcaggca	accacattaa	cacatgcaca	142320
catacacgaa	attttaaaaa	taaaataaat	cttttaaaat	gagctctaga	acagagtgtga	142380
catcagtcta	ggctacacaa	gaccttgtct	caagaagaaa	gaaatgaagg	tttgctgtgg	142440
atggagaggg	agatgcactt	cctattccat	gaagctgcta	ttttggtgat	tatgggtactt	142500
tgcaatttta	tagagagctg	ctattttctt	tcttttaaa	taatgcttgt	tgttttgact	142560
gtaaaagtaa	taaatgttac	tttggaataa	atagagaagt	ataaagagta	aaaaaaaaag	142620
tcataaccag	tgaagtaccg	ttaacatttc	tgcttctatc	cggccagcca	gagtttttcc	142680
tgtgagaatg	tgtttttctt	ttacaaaatt	gggactattg	tgcaacttact	gttttgtagc	142740
ccactctttc	ccttcacgat	ttattgtacc	cattttctca	cagtattaaa	ttttcagctc	142800
caagtaattt	tccatagcta	actctgtgtt	ctgttacaga	gaagaaatgt	acttaattta	142860
agatctaata	ttggcataat	atttggcatt	atgatgctat	aataagcatc	tttctgtata	142920
aatattttta	tgtacagcca	gtgttttgtt	tgggatgaat	tagcacaaga	gtaaaagtgtg	142980
gggtcaagcc	tctgagtgc	gctgaattgt	tctcaggaaa	ggactagtta	acatccattc	143040
tgaagaatg	tgggaatgct	catttcttaa	gcaacactgg	ttattattac	atactattat	143100
tttgttagta	ttatacat	atttaggggg	ataggaaata	tgctcatata	gttcaatttc	143160
aagagttgca	aaaatatatg	gtgtgggttc	ttttcctgtc	ccctagttaa	gtctcacttc	143220
cagccccata	atctaccttg	ttaaatatac	tgtatataag	gcatatgtag	gtgaataaaa	143280
aacagcctag	tatggtggcg	ttaacctcgg	gttacttgag	aactgctctt	gcagagaaca	143340
tgactttggc	tccgaaagcc	ttcctggaat	tccagctcca	agggatgcag	tgccctctggc	143400
atccttgggt	acgtcactca	tgtgcacaca	tacacatttg	gttttttaate	ttaggaactc	143460
caagtgggca	aatgagatgg	ctccatgtat	aaaggcagtt	atgcaaaggt	ctggatgaca	143520
tgagttcagt	cttcagattc	tgcaagataa	caggagagga	ccaacccctg	cgagttgtcc	143580
tctgacctca	gtacacatgt	catggtacgt	gtgtagtatg	cacatgcaca	gaagtcaccag	143640
cactcgggag	gcagaggcag	gaggatctct	gagtttttag	ccagcctggg	ctacaaaacg	143700
atttacagtt	atataaagaa	actctgtttt	gaaaaacaaa	acaggggttg	gggatttagc	143760
tcagtggtag	agcgtttgcc	tagcaagcgc	aaggcccttg	gttcagtcct	caactctgga	143820
agagagagag	agagagggag	agggaaaggg	agagggagag	ggaggagagg	gaaaggaagg	143880
aaggaagaaa	ggaaggaagg	aaggaagaaa	gaaagaaaga	aagaaagaaa	gggaaagaaa	143940
gaaagaaaga	aagagagaaa	gaaagaaagg	gaaagaaaga	aaggggaagg	aagaaagaaa	144000
gaaagacaaa	gcaaagcaaa	actaaataaa	atacatataa	tataataaat	tttaagactt	144060
gagggcatag	atcagtgtta	taatgcttac	ctagcatgcg	taaaactctt	ggcttctaaa	144120
cctagcaccc	taccgcaaaa	tatttgcctc	gtcttgcata	attatattgc	tagttgtcag	144180
actactgtgt	actttcacta	gcaacataat	gagaaatgtt	actatccac	tcctctgtca	144240
agtaactctgt	tcttggtttt	atttttcttt	gccaaattga	tgggtgaaca	agtatttcag	144300
ctagcctaga	acacacagag	aactctcttt	atgaactcaa	gtttcttata	cttttatgat	144360

ctccagaggt	ttgtttttgt	gggtttatta	gtgttttttg	tttggttggt	tgtttgtttg	144420
tttggttggt	tggttggttt	tgctttactt	tttcttattc	atttttttat	ttctttattt	144480
ttttgttttt	aatttaattg	actggttcca	tgtggccaag	gatagcctca	actttgtagc	144540
agaaactggc	tttgaacttc	tggtcttcct	tcatctacct	cccaagtgat	gggatttaagg	144600
cacgtgccac	cacatctaac	aatatctggg	tttctttatt	ggagtttgaa	agggattcct	144660
ccagcattac	tttgactctt	catagtttct	tccagagtta	ttacactttc	atttgttaca	144720
ttaagagttt	gatccagggc	tggagagatg	gctcagtggt	taagagcacc	aactgctctt	144780
ccagaggtcc	tgagttcaat	tcccagcaac	cacatgggtg	ctcacaatcg	tctgtaattg	144840
gatctgatgc	cctcttctgg	tgtgtctgaa	gacagctaca	gtgtaatcat	ataaataaaa	144900
taaatnaattc	tttaaaaaaa	aagagtttga	tccattttaca	ctggacttct	tgaggcgaca	144960
ggatcataaaa	tccaaggtga	gcctgggtga	actggcagaa	gtggcagaag	ctggtctctca	145020
cactgctgta	ttcattttcct	cattgatctt	cagaggtttg	ctaacgggaa	gtaagtggaa	145080
cagaaggttc	agtattcttt	ttttcccaat	tctattcagt	cttttagtagt	agatccctca	145140
ttatctgaga	tgcagagtcc	cctttattcc	tgtgaccatc	tctgtgtttt	tcagggtgtg	145200
tctcattcaa	ggcctaacac	tgaggacatt	tcactgtgct	atgccaatcg	ctctgcagcg	145260
ctcttccatc	tgggtcagta	tgaagtgagt	attgaagaac	ctggtgtcct	gcctgtggct	145320
gcagtggaaa	atgagctcct	ctctgttctt	ctgcacacat	tgaaatcaac	tagcttgcaa	145380
acactggacat	ccaccagac	ccattctctc	ttctgactca	tgtcacctct	cataggtgac	145440
cacaaacaat	atgtagtgtga	caagaagtag	ctatgtcatt	gtccacagtg	catggatttg	145500
ttccaatagg	ccagcacttc	tgtgtccata	tcagctagat	gtgctgctga	tagtatttta	145560
gattccaaaa	tgtgtccaga	tattacctcc	ttcgtttgtt	tcttctaaat	aagccaggca	145620
caaagacttg	aaagatggct	tgatggctct	tccagagggt	gggtttgatt	cccagcccca	145680
acatagcagc	tcacaatagg	ctataacgtc	atttccaaag	ggctctgactt	cctgttctgg	145740
cctctacagg	cagaagcaca	agacatacat	gcaggcaaaa	cacataaaca	taattgaaag	145800
aagatatttaa	ataatagccc	acgcttgagc	ttattcctct	gatgatacag	ctctcctgaa	145860
gtcatcatgg	gcagtgtaaa	agtaaagggtg	ccccgccttg	cccaggggca	tcagtgaggt	145920
agctgactgt	gagtggcttt	cttctcatcc	cctaactgct	gcaacatcat	caactgtgga	145980
gcattatatac	cgtggatttt	aagttaggaa	atgacaaaga	ttagatctat	ggccaggcac	146040
tagtgacgc	ctttaatccc	agcactcagg	aagcagaagt	agggtgatct	ctgtcagttt	146100
gagtttatac	agagtgtcta	ggcagccagg	gctatataga	gaaattctat	ctggaagaa	146160
taaacaaatc	agatctgtat	ttcaggaaga	tgcctatgag	ctgtttgaca	tgtgtgatag	146220
aggctccttaa	ggacaggaaa	gtattccaca	tgtgtgcatt	ttacagaaat	tggttatata	146280
ctggagtttaa	ggactgttgg	aatagttgaa	tgttcacact	cagtgttctt	caaatacaat	146340
agaagaacaa	gaatctatct	gggcatgggtg	atgcacaact	gtattcctaa	catgtagaag	146400
actgaggcat	gctatgtgtt	tgtggctaac	ctgggctaca	tagtcaatat	tggaacagtca	146460
gagctgctac	taaaaccta	aaaacaaaa	tctaataatc	tgggatttta	tattttcctt	146520
tttttaaaaa	aagagtgggc	agaatgtctc	tgattttgtt	cagatggcca	cagaacctag	146580
aaaaactgct	gctgctgctg	ctgctgcata	gcacacagct	aatatttgac	tatacgtata	146640
taaattttgt	tgtatacttt	agctgtgctg	tcaacttttg	aaaaaaagta	tcccagttta	146700
tcatttttaa	ttggcactgt	acagaaatta	acagccatat	tagtctagac	acattaaact	146760
tcatttttcc	atttatacag	aagcaatgta	ctgtatttaa	tattcagctt	tatctacagg	146820
ggtttgatta	cagaaactat	caaagtattc	tctaattgat	gaaaaaagat	tcaagaactt	146880
gactgtatag	ccaaaaggaca	agtggagaaa	aacttaggaa	gaattttccc	tttatccctt	146940
ccctatatattg	atcatctctt	ttacttctaa	taatagtggc	cattttattga	acatacccag	147000
gagttccttt	catcacttta	gatataata	ttatctcatc	ctaaaatgac	ctgttgatga	147060
gtcatctctc	tttcagatga	gaaacaaaga	cattgaaaaa	tctaacttgc	ctgcataaga	147120
tcacacctag	cctcttactc	actccatgaa	tattcctttt	ttttttttcc	tttgagacag	147180
aatctcacta	tgtagtcttg	gttgctctag	aactcaatat	atagaccagg	ctagcctcaa	147240
actcagagag	atctgatagc	ctctgcctgc	cgagtgcctg	ggttaaatgg	atgtgtcacc	147300
aagcccagca	aaattttacct	tcttaatttt	ctaagacgtt	tctcctctta	aaaaatggaa	147360
ctattgagcc	agtcatgggtg	agacaggctt	aatctttaat	ctcagcactt	aggaggcaaa	147420
gacaggccta	tgggttcggg	gcagcctgat	ctatagagag	agttctatgg	gttaaagttt	147480
agggttaaag	ttttgagaca	aaactttgtg	tcaaaaacaa	acaaacaaag	ccagactgct	147540
taataagaca	aatcagacat	aatattataa	acaagtatta	gtgtcactca	attaaaaagt	147600
cactcaggag	gtgagatcaa	tccccagcat	aatgagggga	ggaggggaga	gaaggaaatg	147660
aatgggaggg	gaggaggaa	gaagagagaa	aaaggaagaa	tctcaaagca	gaacacagga	147720
tgtaatttaa	ggcctaagct	ctcgactgaa	gttggtccact	tttaatgacc	cttttcatgc	147780
tcattggtttt	ctgtcttcgg	tacatagtga	agagtggaaa	ccaagggtca	tcctggaatt	147840
tccttttggt	ttcaggcatg	tcttaaagac	atagtggaa	caggtatgca	tgggtatcct	147900
gaaagactgc	agcccaagat	gatggtgctg	aagacagaat	gcctggtgaa	cctggggaga	147960
ctccaggagg	caagacagac	catcagtgat	ctcgaaagca	gcctcactgc	caagccaacc	148020
ctggtgcttt	cctcttacc	gattctgcaa	aggaatgtcc	agcatctgaa	aataaagatc	148080
caagaaaagg	agactctccc	agaacccatc	cctgcagctc	tcaccaatgc	cttcgaggat	148140
atagccctgg	gggaagagaa	cacacagatt	tctggggcct	ccctctctgt	cagcttatgc	148200

acacaccctt	tgaaggccg	ccatctagtt	gccacaaaag	acattctccc	aggagaactg	148260
ctggtgaagg	aagatgcttt	tgtaagtgtc	cttatccag	gagaaatgcc	acgacctcat	148320
cattgccttg	agaacaagt	ggataccaga	gttaccagt	gagacctcta	ctgtcaccga	148380
tgtctgaagc	acactttggc	cacagtacct	tgtggcagct	gcagctatgc	caagtattgc	148440
agccaggaat	gtatgcagca	ggcatgggac	ctctaccata	gcacagagtg	ttctcttggg	148500
gggctgctcc	tcacactcgg	ggctcttctg	catgttgccc	tgagaatgac	tcttttagcc	148560
agatttgaag	atgttgatag	agttgtaagg	atgctttgtg	acgaggttgg	tagcacagac	148620
acctgtttac	ctgaaagcaa	gaatctggtc	aaggcatttg	attacacaag	ctaggggagag	148680
agtgaagaga	agagcaagat	agggtgaaccc	ccaattcctg	gatgcaatgt	caatggaaaag	148740
tatggaagta	attataatgc	tatcttcagc	cttttgcccc	atactgaaaa	gcatagccca	148800
gaacacagat	tcattctgtg	catcagtgtc	tccgcactgt	gcagacaact	caaagctgac	148860
agcgtgcagg	cccaaaccct	aaagtcccct	aagctgaaaag	cagtgacccc	agggctgtgt	148920
gcagatttga	ctgtttgggg	agcagccatg	ctgcgacaca	tgctacagct	gcagtgtaat	148980
gcccaggcaa	taacatccat	atgtcacaca	ggtaagtcat	aaatggtttt	tacttacatt	149040
atttggatatt	caagagctaa	tgtttaagga	gaaaacact	ataaagggaag	gtacggcatca	149100
aataaatcag	tgacctaaaa	ggaaaacaga	ggcgtcttat	atatcattat	gctattgaga	149160
agctttgagc	acatttctgt	gaacccagag	cttgggaggt	ggagatagga	tgattaggag	149220
tctaagacag	cttttagctat	acagcacgtt	tgaggtcagc	ctgaactaca	tgagaacttg	149280
tctcttaaaa	acttgagcca	gagccagggtg	gtgggtggcac	atgcatttaa	ttctagtact	149340
caagaggcaa	aggcaggcag	atccctgaat	ccagcctcgt	ctatatagtg	agatccccac	149400
caggctacat	agtaagatcc	tgtttcaaat	aaataaatat	aacaaaaaca	gcaataataa	149460
caatagcaac	aaattaattt	tttagatgta	tttattttatt	ttatgtatga	gtacaccatt	149520
gcttttttca	gacacaccag	aagagggtcat	tggatcccat	tacagatggt	tgtgagccac	149580
catgtatgtg	gttgcctggga	attgaactca	acacctctgg	aagagcagtc	gggtgctctta	149640
gccactgagc	catctctcca	gtccattaat	taaaaattta	aaactagagt	atttttaaac	149700
atattattcat	tttgtgtgtg	gtatacatat	tataatacag	gttcataagt	caatttctct	149760
ctaccatgtg	tgtcttggag	atcaaaactca	ggttcttagg	catgggagca	agtattttact	149820
tcttgaacca	tctccctagc	catttcttag	attcttttct	tttgtcttga	aagattttatt	149880
tattatatgt	aagtacactg	tagctgcctt	caaataccgg	aaaggggaat	caggcttctgt	149940
tagagatgat	tgtgagtcac	catgtgggtg	ctgggatttg	aactcaggcc	ctccagaaga	150000
gcagtcagtg	ctcttaactg	ctgagctatc	tccagcccca	tttttagtat	tcttattaag	150060
tggtttccat	tttatccaaa	gatgccttta	agggcctggg	aagatggctc	agtgggcatt	150120
gaacttggtg	tgtgagcatg	aagaccagag	ttcagatccc	tagcaccag	gcagatgctg	150180
aatgatggg	gctgcctga	gattccagga	caacggagac	agacaggggc	cctagctaac	150240
catactacac	ctagctgag	ctgtgtgctc	aagagagcag	ccctggctta	ctgtgcagga	150300
tggagagtga	tcattctccac	aggcaagcac	acacctgagc	acacagacat	gcacaaaagga	150360
aagaaaagtc	cttttaagggt	ggtgggtggg	ttgttgtttg	ggggctcttt	gttttgtttt	150420
tttctcccc	tccctcattg	tgatggcaca	ttcctttaat	ctcacatctg	ggacaaaagag	150480
gccggaggat	ctttgtgaac	tggaggtcag	cctgttctac	atagcaagcc	catttcagcc	150540
aggacgacat	agatataccc	tgtctcaaac	agacaaaaat	tattttatttt	atatattttga	150600
atgttttgcc	tgcattgtatg	tctatgcaca	tattgtctgg	tgcccatgaa	agccagaaga	150660
gggcatcaga	tctctcagaa	ctggaatttc	agacacttat	caagtactgc	ctgagtgtca	150720
ggaatcaaac	caaggtcttc	tggaaagagca	gcaagtagtc	tttttttttt	ttaatatttt	150780
tttattacat	attttctcca	attacatttc	caatgctatc	ccaaaagtcc	cccataccct	150840
cccccccccc	ccccgagcag	caagtattct	tcattgctgg	gccatctccc	catctccttt	150900
tctagttaat	taagctgaaa	gggaggagg	tagatgttgc	ccaaacttag	gattttattga	150960
cagattaata	ctctgttagc	ctaactacac	tatagaagct	tattcttttag	actttcacat	151020
tacactgtcc	agattttggc	atcctttttg	ngtgtatatg	tctacagatc	ttaattcagc	151080
tgccaatttta	tacagtgttt	ataggtattc	tttgtgacgt	ggatctttta	cccatcttaa	151140
agcagtagga	tttgaaagct	gacatttatg	tggcctatgg	tccgtttaaa	tcacatttca	151200
agttagtctc	tgtggtagac	attttggggt	ctatctgcgg	ttccgcatct	cacacttttc	151260
cctctcaggg	tgtccagaag	ctgctgcaca	ctgggctgga	aggatgaagt	ggagtccaga	151320
gtgagtggaa	ttctgcagca	tcccgggtcca	gctgggagtg	aatgctgggg	tcaggaggag	151380
atgggtgaga	gggccttctc	caagggcctt	cttagtggtta	cagctctagg	caaaggcctt	151440
cttgacaat	cttagcctgt	gcatagtttt	ttattcgaga	tgagcttgta	tgcatacact	151500
ttattggcag	taaatcagag	gttatccact	cttaggggaag	gagataggaa	taccaaggt	151560
ggacagaggt	cattggctga	aggataacgt	actgagatgc	tcattagcac	ggggaggcat	151620
cccaggaatc	tcagggtgctt	gctgactggg	tttcttaggg	ggttgagagg	gtagcagtga	151680
tttcaccaag	gttatgtatg	gcagagggtta	taggggtttc	aggctcccca	gacaaaagaag	151740
gagaaggaga	agccctgctg	ataagggaag	tccccattt	tgagccactt	cagaaggcta	151800
tcaagacact	gtagaccttt	gtccttaatt	agcagggcc	aacaagtgtc	tgttttcttt	151860
tctgccttca	ttggctcttt	gagccactgc	tacaaaatat	ccaaatctgg	gccaggaagc	151920
tggctcggct	agtaaagggtg	tttgctctta	agcctgaagg	cctgagtttt	cacttgattt	151980
ggtttggttt	ttgttttttt	gagacaaggt	ttctcagcat	agccctgggt	attctggaac	152040

tcaactctgta	gatcaggctc	aacttgaatt	cagagatctg	cctgcttcta	catccccagt	152100
gcttagatta	aagttgtgag	ccaacactgc	ccacctaaaa	aaaatatgag	gggctgggtga	152160
gatggctcag	tgggtaagag	caccgcactg	ctcttccgaa	gggtccgaagt	tcaaatccca	152220
gcaaccacat	ggtggctcac	aaccacctgt	gatgagatct	gatgccctct	tctgggtgcat	152280
ctgaagacag	ctacgggtgt	acttacatat	aataataaat	aaatcttaaa	aaaaaaaaaa	152340
aagaacacta	ttcacaaggt	acaacaacag	tgtactagga	aactttaaaa	tagcccatca	152400
ttcattctag	gggaaaaatc	tttttaaaact	ttagtgtgta	agaaagagag	aggggctgta	152460
gaaaggccac	agcacgtgta	tccagggttag	gagtcaactt	ttcagaagcg	gagtctcccc	152520
ttctacctgt	ttttgaggca	gtctcttggt	tctgccctac	actttgtaca	tgaacttcaa	152580
gatggttgct	ctatttctgc	ctctcatggt	gccctatgca	tgctgagctt	actgatgcca	152640
gccaccacat	aagcatggca	ccagcactga	gccaaaggca	ggcatctggc	tcacatggag	152700
agttacccat	caagccatct	tgctagcccc	agaaaatgta	tttttgacag	gtgtgggtggt	152760
gcacatattt	aatcccagca	ctcaggaggc	agaggcaggc	agatctctgt	gtctgaagcc	152820
agcccagttt	acaaatcaag	tcccagaata	gccaaaggcta	catagagaaa	ccctgttttg	152880
aaaaaacaac	atccttctgt	gttccagtca	caatgactgc	tgtacaata	atatgaggat	152940
ttgggcgtgt	caaaaaatcac	aagtagcaag	gtgtaatggg	caggccttta	gtctcagcac	153000
ttgggaggca	gaggcaggag	gatctctgtg	agtttagcac	agccagggct	gttacacaga	153060
gaaacccctgt	ctcaaaaaaa	ccaagcaaaa	atagaattac	aagttaacca	gggtattggt	153120
gttaatatga	aatagcagaa	ctcaagatag	ctaataaac	aaaggattct	attttataaa	153180
ggagacattc	catactgaaa	tatatgcaga	gcctgatgct	tgcctagcaa	ctgaactaca	153240
cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153360
aaacgtcagg	ctcagaggca	ggatggcagt	gatgtgtttt	ggggcacaag	gccccctttc	153420
attaaagcac	caaatcttgt	actaaaaaca	gccctctgct	ctcgatctgc	agagatacga	153480
gagaggcaca	catagcttcc	gggcccttgc	cctctgctct	cccggtcagt	tcctggactc	153540
ctgggagaag	cttgaagctc	aagaactggc	cgctgttgcc	tttgtgctga	cgccagaacc	153600
agcagctatt	cagcagctgc	ggctcttggg	caagcttgga	agtagctcgt	tccttctcct	153660
tctgcagggg	aaagacaagg	agatggcact	gagctgggta	gccaggatgt	gggaagtaaa	153720
ttgtgtctat	gtgggtgagg	gaagtgccgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	153780
gtgtgtgtgt	gtgtgacaga	gacagaaagg	gagagtgcac	gttcttgtgt	gttctgttga	153840
gtcagtgtct	cctttggctc	ctcctgccaa	aaagcatgct	gtgttctcag	agtcaagctt	153900
ggccaggctc	gcccacccca	ccaaagccca	ggatctgcca	ccccattaag	aatgcggggg	153960
taggaaataa	aaaacagggt	tcttgtccaa	gagaattttt	attattattt	tttctctctt	154020
aattggattc	agcatttctt	actcctcagt	atcctctctg	gtaggggaatt	cagggtctgtc	154080
tatgcagagc	acaagagact	gtgcttgcca	gaaacccctg	ggccaacagc	cctattccct	154140
gactggctct	ggctgcaggc	tgcccttctg	gctgacccct	gagtctggcc	ctctgacctc	154200
tgcccgctct	ggggtcaccc	aggaggagca	ggaccactgt	gatacagggt	tcctgagatg	154260
gctactgaaa	caccctcatc	tgctaaggcc	actagatttt	tttccactg	ccagactact	154320
caggggccctc	agtaggtcac	tggccaaaga	gcctagatgt	taaaactgca	gctaaagcct	154380
ctcctgaggc	cagagctcag	agcctccctg	gcctgcacaa	agtgtcaaga	aggacattgt	154440
ccatccaacc	cattggacta	acactgtaga	cgctgccttg	tctgccagct	tgaaaacccc	154500
agaaacctct	tccccagctc	cgcctagctt	gctccagca	ccctgacatc	cacttctctc	154560
gctaagtctc	gcagcttcta	cagtaggggt	gacgggggtg	tcgggggtgc	cagacaggcg	154620
tgtgttgcac	ataaaaaacga	ggtgatgttc	taagtatcta	agaatgttgg	tcccctgaag	154680
tgattcttgc	tgcttctctt	ccttcccagc	tcttccact	gacacctttg	ccccagcaa	154740
gcccaggggac	tattgttgct	ggctgggggt	ctgaacaaaa	ttgccgcagt	tttttgtttt	154800
tttttttctt	gctagaagtt	accgatacag	tccttaattg	agcaaatata	ggttcccgtg	154860
tagtttataa	acatgataag	acatacattt	tggttaaggag	tgggttgggg	gacctgtgca	154920
tttatatat	tatatatata	tatatattat	gtgggtgtgt	gggcagagt	aggatatata	154980
taagtggaca	taggtataaa	actgcacctt	ctgtgtgact	ctcattgcga	gtacagttct	155040
aaatgtcatc	cactggcgat	ctctccttgg	gtgattgggt	cttggacccc	agcaggtctg	155100
ccgggggctg	cctgagacgt	caggaatgag	aggcattacc	cccatggata	gggactgagg	155160
gtggcatagg	gttggacagg	gcagggttaac	taagtgatct	cagacaagag	ccaagagtgc	155220
tctgagattc	ctgggtgccc	cagctggctc	tgggagagcc	ttgttctgag	tcctgctcct	155280
tccaaaacca	gcagggtcct	tcagcccttc	tctccaaatg	accaggcttc	cgcagagccc	155340
agcttcttca	aggggcgcac	gtcccgcac	cactaatgac	tcactttgcg	tgcccttgac	155400
cactgtgctg	gagtggatac	ggtccagagg	cgctcgggtc	ggacagccga	gtgagacgtg	155460
atacccttcc	cgtctacggc	tgtacatttt	gggcttataa	tccaccagga	agggctccaga	155520
cggtggctga	aggcttcagc	agccttttcc	tgaacccag	cagatcttcc	acttaggaaa	155580
aaaaaaagaa	agaaagaaag	aaaagaagaa	aaaaattctg	ttccttgctg	gacatgtggc	155640
agtgtctggg	gacggagccc	tgggtccctc	gcggagagt	actgccagcc	ccagtatcca	155700
ggccaggagt	ggggcccaaa	gggcccgtgc	tggagatgcc	tgctgcaccc	cactgggggc	155760
acggatgggg	gtcctgcgag	cacacttgcc	cttctctctg	ggctcgtcgg	tgggcatgat	155820
gtcaaagctg	aacttgtgct	gatagtcggg	ggcatagctc	tgcagttcgg	taagctcttt	155880

cccagagctc	accttagaga	tctggttccg	gttcctgtgg	ctggtgcagt	tcttgccctgc	155940
cttcttgtaa	cctgacctgg	agccaggcgg	atggccatgt	gggtggcctt	tgtccctgga	156000
ggccccatgg	gacggatggg	gctccttgcg	ggcagccctg	tcagaggtgg	taagcgtgtg	156060
agacttgatc	tggtgaggag	acactggctc	tgtgcagttc	cgaagtcct	ccacctcag	156120
cagcttcaga	tcctggcctt	gccgcagctc	gggggtcgcg	caggggacag	cagagctaga	156180
gccacggaac	cttcgcagcc	attcccacag	ggaacgtgcc	cggcagccac	agtcccaagc	156240
attcccattg	aggcgaagga	actccaaggc	caccaggggg	gccagacagt	cacctgcag	156300
ctcagtgagg	ctgttggtga	agagaaagag	ggtgggttagc	ctgtggaggt	catggaaagc	156360
cttggtgtga	accactgtga	gctggttctc	atgcagcagc	aaccggtcca	ggttcaccag	156420
gccccggaag	atgccttggc	ccaggctcca	tagcttggtta	ccatggagaa	acaagtgaat	156480
gagattgacc	aggtccacaa	agatgtcacc	ttggagggtac	tcgatatggt	tgtcctgcaa	156540
gtagagatac	tgacacacac	gcaggccacc	aaaggtgcct	gcgggcaggg	cgctcagctc	156600
acacttatag	aggtagaggg	cgtgaagctt	caccaggcct	tggaaggtct	cgggtgccag	156660
cgttcgcagc	tgtcgggtgt	ctccaaggtc	tagctcctcc	agatgcacaa	agccctcgaa	156720
ggtgttggga	gcaatgaaag	tgatgttggt	ggagtagatc	cagaggggtga	ccatggcggg	156780
gctgaagtgg	ccctgctgga	ggaaggtgat	gcgattgttc	tgcaaggaaga	tgcgctcact	156840
gtcctctggg	atgcctctcg	ggatggcagc	aaagttgtgt	gcctggcagc	tgacagtcac	156900
gggcgcagtc	tagcacacac	agtctcgagg	acaaccacca	cccagaggta	gctctccagc	156960
gagcagcaac	agcagcaatt	ccacacagga	ccctgggtggg	gagagacaga	acagcagtga	157020
ggggctgccc	agaggaggtg	gagatagatg	gggaacagag	ggtggaatgg	gggactggca	157080
aatgactctg	ttggctcaca	gaggttctgt	cctctgtatt	gtatgcaggg	gtcccttgga	157140
cagggcatatt	ggggccaagg	cccacattat	ctcctcacct	cttttagctct	gtccctaaag	157200
tctctaattc	catccgaaca	cttcttcaga	ctgtcagccc	caccgcaggg	tggaaacacac	157260
ttgtgaacac	agcagagtcg	gcctccgggt	ctgggtccgg	ctctgccact	cgctcactgt	157320
tagctgcctt	atcaaggaat	gactctaaca	aagcaaatct	ggagtccctga	atgatcactt	157380
tatttaaata	attcctcaaa	ataaagaaag	cattgagtc	atggtaccaa	agcatgcctc	157440
aataagcgcc	tctttcacac	tgtggtacaa	aaacttcaaa	cctacaactc	ctccatggct	157500
gtttcctcat	gagttaaaca	cagttcacag	ggctgtgtgt	acagacaagg	cacatttctg	157560
tgaggggctg	tgagtgcac	cagggctgac	gcacgaggct	tccttggggg	ttcacagtac	157620
tgccagatcg	aggctgcatg	ctcctctccc	ccattcacac	ccccccctcc	tgtgctggag	157680
attgccagcg	tgtggctgta	aaaccggg	ttgcctcttg	actgtccaga	gcatcttctc	157740
tgtagcttcc	ctctaattgg	gcattaatta	ggcattcggt	aatggatcct	taaaataatt	157800
attttcggat	gtgtccagcc	tgtggtcggg	taataggcct	atgctcatta	tgggaagccg	157860
ctcattatgg	acgattgtca	ttacctgcct	ttttccaggg	tcacagctgc	cccaagtggc	157920
ccagcaggcg	cgtcaggaag	atggggacag	gctccaggcc	tacgggcg	cacctgaag	157980
ggccaggcag	ccacgaccta	tgtgcctca	ggtggcctct	tgcccttct	tttccagctt	158040
gttcagctgg	gactcttggg	agagccaggg	ccccggggg	aatatgagct	gagctgaatc	158100
ttcttgctgc	tagctgtgct	cagagcaagt	gggaaggaca	gggaccttct	gacaggctt	158160
cccacttggg	gtcccaggcc	cagggactgc	ccaggccccg	gcagagtagg	ttgccacctt	158220
gacttctgac	gccccccccc	cattcccaac	agaaacagca	tcgtaagttg	acagctccca	158280
gctgttggga	gttatgggct	cccagagagt	ggcagctgct	tctcgccct	gtaatcacc	158340
ggcttcagct	agaatgtttc	tagcacataa	aatcatcg	atataattta	gtttttgcat	158400
aattgggttc	agttgtgatt	tcagagcaat	tatgacctca	gcagcagggg	tggcacagca	158460
taggacccct	ttctgggcgg	ggccatgccc	tcgagccct	ctgcagtgtc	ttctgccac	158520
cgcccttag	acagcatgca	ggctataacc	atctctgcct	caatttctct	ctccaaaacg	158580
tgtctagatg	ttactctgtc	gatcttctct	ctcagcatcc	tgggtgtggc	ctccagcctg	158640
ccagcctctg	tcctagggat	cctgtgctgc	agagggaggc	acagtcggag	ggaggggagc	158700
ctgccctgtg	ccaccagcac	tcacactggc	tgacagctcc	acagacccac	agctccaaac	158760
tcctgtcttg	gcttgacccc	tctcttccag	gaaggccatt	cttgccagaa	cctttcccaa	158820
cggtcacctg	ggaaagcctg	gactctaggt	tcaaggacat	tcagtatgct	tgccccacat	158880
tttatgctgg	atgagacaca	gcagagcctt	cttcaactgg	gggtcctgtg	aaaatgaaag	158940
cttttcttcc	ccggcctgca	gctgcaggca	ggtaggggtt	gcagtgggct	tatcactaat	159000
accattcgac	atttgtacag	ctcatcgag	tttacagagg	gcttttgttg	taccctcaac	159060
ttccccctgtg	ttctcccacc	tactgtggct	gctctgtctc	tgtgcacccc	aaaagaatct	159120
ggaagtccct	tggggagatt	acccccctta	catagggggc	tccaaggaat	acaggctaca	159180
gctctatata	ggaaaaaaa	aatcaaacctg	aaccaaacc	caggtgtgga	cttagtaacc	159240
agtttatata	cataccgtgg	caagtggagg	aggcaggcgg	cagaacggca	tgaaggtagg	159300
actgggggtt	tcctctctaa	agggcaatgg	ggggacaaa	ggactctagc	caagacctga	159360
tgacagaacca	ggctcagttc	ccctgttatc	tcaaggctat	catcactagg	aggcctaagg	159420
caatggacca	caaggacctt	gtccttgtag	gacagtcact	tcctgcagtg	aagtgtctt	159480
ctggaagcat	actaatagga	tgtaggctca	ggacagctgg	tctctgtcct	ttagtatatt	159540
tcacagtgcc	agggatgtta	accttccaa	cctccatctc	ttctaattggg	gggggggtgt	159600
tgaggggtc	agccactctg	catccatgtc	tttgaaagcc	agtgggtatta	ctccaggacc	159660
ctgagcaagc	tgtcctagtc	agccctggcc	acttctggac	tccttgccctg	agtcagtagg	159720

tgccaatcct	aggattgtca	ccagcaggtt	tcttccctag	gaggcaagca	ctgtatcacc	159780
atggcgccct	ctatgcccc	tctatgaggg	ccttggggag	cccgccccac	tgattgcctg	159840
attaatgtac	caacaatgag	gatggagcct	ttgccatgca	ttttaacatt	gcaaattagc	159900
aggaatccaa	gtctctgtgg	aggggcccctg	cacctcttct	gccagactca	tcaagcgccct	159960
cttgggcagg	gctgccttct	acttgagggg	gcggaaggga	gaagacccag	ttccactctc	160020
cttccccctcc	aggaggtgcc	cttcatcgctg	ttctgcttcg	ttactctcaa	gcctccggcc	160080
tcccacgcac	gtgagctccc	aaggggctct	acagcctccg	tcattccttc	ttccattcat	160140
acttgcccc	tagtctagga	gagccatgga	agacagtgtg	ggaagggtct	gacaatgagc	160200
atcatgcccc	atgtgcatat	gcggtggcaa	taccctgggtg	ggtaccagga	gagtagagg	160260
gaaattaaga	yaggggccta	aggaaagcct	ctgctatccc	tgggctacca	gtcagcattg	160320
cttggtcact	gatccccctct	gtaacaccag	cccttctgca	acctgccaga	gtttttgacc	160380
tttgaactag	ggctgagaag	ggctctgctct	gttcagctgc	cttggctggg	aggggaatct	160440
gctcagacct	cagcacacac	tcaacagaag	gcatgcaagc	aaggagagcta	gcagtggcct	160500
tgggtcagct	ggcaagcccc	aaactcttcc	tgccaagctg	agcatgaaaa	gccacctcac	160560
catgggtccca	tgggaccaga	cctggtagga	taggtggcaa	ggctaaggca	gcggaatagc	160620
atgtgcaaa	gcactgggggt	gggaaagggc	ctgtgcttct	cacccctctc	aatggtgcag	160680
agcctccaag	gaatactgta	acctcagctc	agctggggtc	gggtggccag	agagcttggc	160740
accagaacca	gcatcaacag	ggcctgtctg	ctaaacccag	acctcacaag	ccagttagt	160800
aggggcccctg	tagcaccctg	gccaccagaa	ctaaccagga	agatctgacg	ctgggaatat	160860
gtctttaatg	aaaagccctt	ccggaagcca	catttgacaca	gaagaaaatg	aggtgccag	160920
agcatcagtg	ggctggttgc	agctggagaa	cacagcaggg	ggacagggtcc	taccaagcta	160980
ccctgcccctc	aggctggggc	tctagccagc	tccctgatgc	ctggagtagg	taaagcagcc	161040
tcgaaatggg	ctgggtcagc	tttttcaggc	tccaaagggg	caggacagct	gtcagcgtt	161100
agcaccacaag	ggggctgccc	ctctacccct	aagtagaggc	atccccatgg	cccctgggca	161160
ggctcagtggg	tctctctgaa	gctttgtagg	ctgcttcttg	gccatgtagc	caatctctct	161220
ggccttcagt	cctcccctccc	tgcccccagc	ctggccagct	gctcttctct	gagcaatcga	161280
tgtaaaccca	atgctctctt	gctgtgggga	tggcggcctc	aggccaggcc	agctgcactc	161340
ctggggctgc	tggcgccctca	ggccacttgg	cacttgtgcc	acttgtgttc	taaacacagg	161400
ctccttccctg	gctcggccct	gaagacaaga	aagctggcca	gggaacagct	gggctcccat	161460
ctcagctccc	actgctgtgc	agagcggccg	gcagcctcct	atccatggct	gtgagtagaa	161520
cagggctgtg	gagccagagg	cctaagttga	atcctggctg	ctccttttaa	tgcttgagg	161580
agcctcgttt	tcctcacctg	caaaatgggg	cagtcattgg	aagctcagcc	agtcctccag	161640
cccacagata	ctggtaccca	cctgcccttc	ccacatctcc	atctgtctaa	ctgaaaccat	161700
ccaaaccaag	ctcttctctt	cctctccggc	ctcctccgtt	cacaacttct	ccatcttggg	161760
taaagatggt	ccctttacat	cagctgcctg	ggaggaacac	ctcagaatca	ccgtggctca	161820
ctctggggaa	atctgtctgg	ctagatttta	gaatgtatcc	agtatctatc	ctgttatcat	161880
actcgctatt	gctaccattc	accagtagc	ctcctggatg	ccctccccc	ccccccgcaa	161940
gccctgcctc	ctcaccctta	cacctccttc	aacagggaact	agggtagtcc	agggaaagtg	162000
agtcaggaag	ggctgctcct	tagtctgcat	cctccagagt	gcccattctaa	ctagaagctg	162060
cccaggtct	ttctcccagc	ccagaggccc	tgcctctccc	tggctacaca	gaccttgctg	162120
ccgtcccctcc	cacatgccag	ccctcagcct	ccctcctgcc	tttgtccatg	ctgttccatc	162180
tacctggacc	ggttcccagt	gtgtctgcag	ggctgcttcc	cagagaagcc	actcttgagc	162240
agcgatttgc	aaggagcttc	tttccctggga	cattttactc	tacagcacgc	gacctcttgg	162300
acagcacgac	ataagtcact	tgtttccctt	atgtccgctg	ccactttgtt	ctgatgagtg	162360
tgctccctgc	acacagtagg	tgtcatttaa	cagttctggg	ggaggggaatt	accttctcaa	162420
gtctctggag	aattgaatga	tgacactcag	gaagccctag	gctcaagcct	ggggcctgga	162480
tcatagtagg	tgctcgataa	atggttggtg	taattagtcc	tgggagactc	agagccttca	162540
ggagaacaga	cacctgaact	tggctcacat	caagactcct	taggcccac	agggaagtga	162600
ccgtttgttg	gatgagcact	ctgagaagga	cccaatacca	gtcctttcct	gggcaagggg	162660
aaatagactg	ggactgggga	gtttcccaag	gtatgggatt	ttcagcacta	aactggaata	162720
atgctaaaga	aaaaaaaaaag	ttagtcactc	taaaaggggc	caggacccaaa	acctttcaaa	162780
cagaaatgtc	tgggtttatg	aagagaggaa	gccagatatg	gtggggcaca	tctttaatcc	162840
aggtgcttgg	gaggcaaaga	cagatggagc	tctgtgagtt	tgaggccagc	ctattctaca	162900
aagtgagttc	taggacagcc	aaggctacat	agagaaaccc	acttgacttg	ccaccccaat	162960
taaaaatcctt	agtgggggag	agtccaagga	ggaaacacac	acacacacac	acacacacac	163020
acacacacac	acgttagtat	aatatcatac	tatggctctg	tgcttgagct	ccaggaatga	163080
gggctgaact	cagagtgtta	gtgtgtgcta	gtggatattt	gagctctgta	tttatgtgca	163140
tgtctgtgta	gatgtgtacc	tgaggtgttt	atgtgtacac	aggtgttggc	ctgttgcata	163200
tggatggaga	catggttgtg	tttgctaggc	atgtgcatgt	gtctgagttc	atgcacataa	163260
actcacatct	acctctggag	actgagagtg	acaacccagc	gcccccttat	ctcgcagcac	163320
cccaggccca	gcaccccagc	ccagcatccc	aggcccagca	ctccagaccc	agcaccagc	163380
gcccagcatc	ccacgcccag	catcccaggc	ccagcaccac	agaccagca	tcccaggccc	163440
agcacccccag	acccagcacc	ccaggcccag	catcccaggc	ccagcaccac	gaccagcat	163500
cccaggccca	gcaccccagg	tccaagcact	caatgcccag	caccccgacc	cagcatccca	163560

ggcccagcac	cccaggccca	gcattccaagg	cccagcatcc	cagggacage	accccaggcc	163620
cagcatccca	ggcccagcat	cccaggggaca	gcaccccagg	cccagcatcc	caggtccagc	163680
atcccaggga	cagcacccca	ggcccagtat	cccagggaca	gcaccccagg	cccagtatcc	163740
cagggacagc	accccaggcc	cagtatccca	gggacagcac	cccaggcaca	gtatcccaca	163800
tggaggcagc	acatactgaa	gatagggaat	gtctctgagg	cctcttatct	tggctccttac	163860
cctcattgct	ttcagcacct	gctctcctca	cactcggaat	caaacaccct	gtgcagggttc	163920
tcccagtagc	aggattcccc	tcagctgagg	aatgggtagc	taccattttg	gcttttgtct	163980
gtctgggggt	ggcagcccca	tgctaattgg	actgacagtt	tctcctgaga	gcaatttggg	164040
cagcacatcc	tgcccattag	gcctaaccct	gcctgcaggg	gtgtgctgta	ggggcaggga	164100
tggagcctac	cctgtatagc	tctgtattga	ggcactcccc	caagctatga	cccattgccag	164160
tgggagtcac	ttcacctagg	caactccaga	tgggcacaaa	aatctctcca	ataagggtag	164220
gtatgggaat	aggtaaggag	agcatagtga	gcctggctgg	gcacctgaga	cctgagcagc	164280
ctgcacggga	gatttgtgtca	ctgtggttcc	agactgccaa	gacatcttgg	ctttcacccc	164340
aactcaggat	ggtccagaat	ccagagctct	taagagagca	gatgctgaga	ggcacttaac	164400
ccagggtcaa	gacccttccct	tggacgggtt	cttggctttc	tactctgtcc	tctgtcccag	164460
tctgtcatcc	ccatctgtgc	ctaacagctc	tctgtggaaa	acatgaggcg	tatgagctct	164520
ctacttctcc	cagcatccca	tgcccgcacc	ccagctcact	gtgtgccctc	atgttactca	164580
aatcttctgc	taggtttgag	ggcccagggt	tgaggctgtg	ggtttccctc	catctgtccc	164640
tccctttttac	caccaccact	aatcctcttc	ctcctcttcc	tcctcttccct	cctcctcctc	164700
tttctnctcn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164820
aaattgaact	tacagaggaa	acacaaatga	gctgtaatga	cagaaaagga	atactatttc	164880
ccttgtgtac	agaaagctat	agattaaacc	ccagcccttt	ttgcctttgt	ttttcatatt	164940
tagtgtttct	ggggactgaa	cccaaggagc	gcagctgctc	tggaaatgtg	gagacctctc	165000
aggcattgct	tatgctaagc	tgacaaaggc	tttctgaatg	catatggtag	tatctattca	165060
tcttaagatg	cctataacca	tgggtggagc	aattctctaa	gagtctgtct	gaacttctta	165120
agcatgtgct	caaaaacaca	cataaagatt	gtttactgaa	gcagtatttg	caacagtga	165180
acaaaaccaa	acaacaacaa	aaagtactta	atagctcatc	agtaggaaaa	agaacaacaa	165240
attgggacac	atttatacta	ttctcgttct	atcacacgac	tattaagaag	attaagggtg	165300
gaggatgaaa	agcccctgca	gaacaacgca	cacaatgaaa	ttgaatagag	aataaagaaa	165360
aaaatatata	gtgtatttat	gtaaaacata	cataaacaca	taaaaagaat	gaaaaactct	165420
ttgactgtat	atctttgaaa	agaaccagca	gagtgtctatg	gtttataaat	gaaaagttcc	165480
cccaaaattc	ctgagttgga	ggttcaatgc	agagttcata	ctcttaggaa	gtaactaaat	165540
caggaatgct	ctgatcta	caatggattg	accaactgat	agattaataa	tctgaaggca	165600
acacttccag	gaggcagaaa	caagggtgagg	cctaactgga	ggaagtttgt	caccagaggc	165660
atgtcctttg	agaggctacc	tcactcctga	tccttccctc	tctccgcttc	ctggctgtaa	165720
aagtatttat	ttgactgatg	taactgtcat	cttggaggct	tactggctcc	atcagctaac	165780
ctaggcctag	ccctggaagc	ttctagcttc	catacaatct	aatccaagcc	tagaatgttc	165840
cagcctttag	gacttgctgc	tgagatcacc	gtttcctgtt	ctttctgaac	tctagctggc	165900
tgattcagtc	cccctgttcc	gggctcaaac	tcctctcccc	gatgatttta	ttcacaatct	165960
gtcttttctc	ttggcctctg	aattgctctg	cttgggtctca	aactaactct	agcaatcttt	166020
tctaattctct	tgtctccttc	acactctctt	gcttgttctg	tctttactgt	gtctagtttg	166080
ttctctctct	catccttctc	tgtaaagctc	tcccggtaaa	cctgcctcct	cctccccctc	166140
tgtgccgctc	tactctccct	ctcagctcta	ctgcactgct	ctccacagct	ctcctgtatc	166200
ctgtgctgca	ctctcttctc	cggtaaccacc	tgtgtctccc	ttacgtagct	tccctttcct	166260
ctctcttctc	ctgagggttg	ggcagatcct	atcctgtcaa	acctttctct	gattcttcac	166320
tttgtctgcc	actcaattag	acatcacttt	caagcaggag	tgccctcctct	acaaaccaac	166380
tttaccttca	ttgtttcaaa	ttaaagggtga	gtactaaggg	tgtgtctctt	tttcagccag	166440
tgagagtaaa	gatgtgtgct	aataaggctg	agccaactct	agctagaaat	agttcttttt	166500
tctccataaa	taacagaatc	ttagggttca	caatacgtac	aaatatcctg	agacagctgg	166560
ctgccatgtg	gtgaagagaa	tgacagagag	taaggatgtg	ctcttgagg	ttcagatggg	166620
aataagcatt	ttcttgggag	ttggattaga	gtcattcctg	tgacactgtg	acaaaggact	166680
cgactacatt	ttgccatgcc	ttgagactgt	ggagggtgga	gcgatggatt	aatacactgg	166740
agtgaatttt	aaggcagcca	aggctgtggt	tggactgtta	ctagctacgt	ttagctggat	166800
ttatatatta	aattgggaac	aaaaaagcag	agtagaaggg	acagtttttg	cagaaggagg	166860
tttcttcata	tacaacttag	tttttatagt	tagcttttta	tgttgctatg	accaaataac	166920
cttatggaaa	ctgaagaata	aaaattttta	ctgaactctt	cttcacccca	gaacccgacc	166980
cctcccatct	agagattgtt	cccgaacac	tcctgaactc	ttcaccccag	aatgctttcc	167040
tgaactctc	accctagagt	tcgaaccctc	ccaactaaaa	actgttccaa	gaacattttt	167100
gagataaggg	cctcctaata	caacctcaaa	atgaaccggg	tacattgcca	aataatagga	167160
catgacccct	tagttacgta	gattcccttg	gcagaccccc	ttgtcccttg	acagaacccc	167220
ctagtgtagt	aaacttgtag	tttccctgcc	cagctctccc	cccttgaggt	ttactatata	167280
agcctatgaa	aaatttggct	ggtcgtcgat	tctcctctac	accactaggt	gcatgagttt	167340
cgacccca	gctctggtct	atgttccatg	tgttttcttg	ctgttgttct	attaaatctt	167400

gccttctaca	ttttgagtag	ggtctcagtg	tcttcttggg	tccgcggctg	tcccggggct	167460
tgagtgttg	agttaggggc	tcccttcggg	ggtctttcat	tttgggtgat	tgcccgggaa	167520
acagcgcgac	cacccagagg	tcctagaccc	acttagaggt	aaggttcttt	gttctgtttt	167580
ggtcagtgag	accgcgctcc	gagagggaa	ggttgcgac	gcagtttcgg	ttttgcggat	167640
tggtccacgt	ccgttcaccc	tgggagacgt	cccaggaaaa	acaggggagg	accagggacg	167700
cctgggtggac	ccctttggag	gccaagagac	cattttgggt	tgcgagatcg	tggttttgag	167760
tcccacctcg	tgcccagttg	cgagatcgtg	ggttcgagtc	ccacctcgcg	ttttgttgcg	167820
agaccgtggg	ttcaagtccc	acctcgcggt	tggtcacgag	atcggtgggt	cgagtcccac	167880
ctcgtgcaga	gggtctcaat	cgcccggtct	tagaaaaggcc	atctgattct	ttgagttgct	167940
tgtggtcgag	gcagagtcgc	cgccgtttct	ggttcttttt	ttgtcttagt	ctcgttcgct	168000
ctcttgttgt	gtctactgtt	tttctagaaa	tgggacaatc	tgtgtccact	cccctttctc	168060
tgactctgga	gcattggaag	gaggtgcggg	tcagagccca	caaccagtcg	gtggaagtca	168120
gaaaggggtcc	gtggcagacc	ttttgcgcct	ccgagtggcc	aacgtttaga	gtaggctggc	168180
cacctgaggg	tgcttttgac	ttgtcactaa	tcgctgccgt	caggcggaat	gtttttcagg	168240
aggaaggggg	tcaccctgat	cagatcccct	acattgtgac	ctggcagaat	ctcgtccaat	168300
tcccacctcc	gtgggtcaag	ccttggaccc	caaactcttc	gaaactgacg	gtcgcgggtg	168360
cccagtcgag	tgacgcccga	aagtctggcc	catcagcacc	ccccaaagatc	tatccagaga	168420
ttgacgacct	cctctgggata	gactcccaac	ctccccctta	ccccctgccc	caacagccac	168480
ctgcagctgc	cccaccacag	ggaccaatag	cgagaggggc	tcagggaccg	gcgggggaga	168540
ctcggagtcg	ccgagggccga	agccccgggg	aggaaggggg	gccagactca	acagttgcct	168600
tgccactcag	agcacatgtg	agagggccag	caccaggacc	taatgatctc	attcctttac	168660
agtactggcc	tttttctctc	tctgatttat	ataattgaaa	aactaaccac	cctcccttct	168720
cagagaaccc	ctctggactt	actgggctcc	ttgagtcact	tatgttctcc	catcaaccct	168780
cttgggatga	ttgtcagcag	cttttgcagg	ttctttttac	aacagaagaa	agaaaaagaa	168840
tcctcataga	ggcgagaaaa	aatgttcttg	gagaggacgg	cacacccact	gccctcccta	168900
acctcgtgga	cgaggctttc	cccttgaacc	gccccaaactg	ggactacaac	accgcggaag	168960
gtagggggacg	cctccttgtc	tatcgccgga	ctctagtggc	aggtctcaga	ggagccgcta	169020
gacggccccc	caatttggct	aaggttaagag	aggtcttgca	ggggcgagact	gaaccaccct	169080
cagtccttct	tgagcgtcta	atggaggcat	ataggagata	cacccctttt	gaccccttgt	169140
cagaggggca	gagagccgct	gtagccatgg	ccttcattgg	tcagtcgct	cccagacatta	169200
agaaaaagct	gcaaaaggctg	gaggggctcc	aagatcatac	gctccaagat	ttagtataag	169260
aagcagaaaa	agtctatcat	aagagggaaa	cagaagaaga	gaggcaggag	agagagaaga	169320
aagaaataga	ggagagggaa	aatagacggg	atcgccgtca	ggagagaaat	ctgagtaaaa	169380
ttttggccgc	agttgtgaat	gatagacagt	caggaaaaagg	taaaaataggg	ctcctgggca	169440
acagggcagat	gaaaccgcaa	ggtggcaaaa	agataccact	ggaaaaagac	caatgcgcct	169500
attgcaaaga	gaaaggacac	tgggctagag	attgccttaa	aaagcgggag	cgatccaagg	169560
tcctaaccct	agaagatgat	tagggaagtc	ggggctcaga	ccccctccct	gagcctaggg	169620
taactttgtc	cgtggagggg	actcccgtca	acttccctgat	agacaccgga	gcagaacatt	169680
cagtactcac	taacccccta	ggcaagctag	gtcccaaaaa	gaccatggta	attggagcca	169740
ctggtagtaa	attttaccct	tggacgacca	aacgagctct	tcagatagac	aaaaatatag	169800
tgacccactc	ctttctgggtg	atacctgagt	gcccctgctcc	cctcttgggg	cgcgatctgc	169860
taacccaact	aaaggctcaa	gtccaattta	cttcagaagg	cccacaagta	agctggggaa	169920
agggccctgt	tgctgcctt	gtcctcaaca	cagaaaaaga	gtaccggttg	catgaagaac	169980
aacccaaaaa	tgagctctct	tcaggttggc	taactgcgtt	ccccaatgtc	tgggcagaac	170040
aagcaggaat	ggggttggct	aaacaagtgc	ctccggttgt	ggtagaactt	aaagctgatg	170100
ccacccccat	ttcggtaaaa	caatacccca	tgagcaaggga	agctagaaaa	ggcatccggc	170160
ctcatatcca	gaggttgctg	ggccaaggag	tttttagtggc	ctgtcagtc	ccctggaata	170220
caccactttc	gccggttcaa	aaaccaggga	ccaatgacta	tcgcccggta	caagacctcc	170280
gggaggttaa	caaaaagggtc	ctggacattc	accccacagt	cccgaacccg	tacaatttat	170340
taagctctct	cccacctgag	agaacatggg	atacagtcct	agacttaaaa	gatgccttct	170400
tttgccctgcg	tttgaccctt	aagagtcagc	tcctgtttgc	ttttaaatgg	agggaccacg	170460
agggcggaca	gactggtcaa	ctaacttgga	ctaggctacc	acaggggttc	aaaaattccc	170520
ccaccctgtt	tgacgagggc	ctccatcggg	atcttgcgcc	ttttcgcgct	cgaaaccctc	170580
agcttaccct	actacagtat	gtagatgac	tcttggctgc	ggcgccctcg	aaggagctgt	170640
gtcaccaggg	aactgagagg	ctcctcacag	aactgagtga	cttggggtat	cgagtttcgg	170700
ctaaaaaggc	acaaatctgt	caaactgagg	taaccttctt	gggtatatac	ctccgagggg	170760
gcaaaaagatg	gtcacagag	gcccggaaaa	agactgttat	gatgatccca	tcgccaacta	170820
ccccacggca	ggtacgtgag	tttctgggga	ctgctggctt	ttgtagactc	tggattccag	170880
gctttgcaac	cctagcagca	cctctatatc	ctttgactaa	ggaaggggtt	cctttcaagt	170940
ggaaaagaaga	acaccaaaaga	gcttttgagg	ctatcaagtc	gtctctaagt	actgccccca	171000
cgctagcatt	accagacttg	actaagcctt	tcgtctata	tgtggacgag	agagcgggtg	171060
tagccagggg	agtattgaca	caagcactgg	gacccctgaa	aagacctgta	gcctatttgt	171120
caaaaaaatt	agatcctggt	gctagtggat	ggccacatg	tctgaaagct	attgcagcag	171180
						171240

tagccctgct	gatcaaagat	gctgacaaac	tgacaatggg	acagcaggtg	accgtttag	171300
cccctcatgc	cttagaaagt	atcgtgcgac	agccacctga	cagataagat	gacaaatgcc	171360
cgaatgacac	actatcagag	cctgctgcta	aatgagcgtg	taacctttgc	gccccctgcc	171420
atcctcaacc	cagctaccct	tctccctcta	acaaatgatt	ccgtcccagt	acatcaatgt	171480
atggacatcc	tcgctgaaga	aactgggacc	agaagtgacc	tgactgacca	accctggcct	171540
agagctccca	gttggtacac	ggacggcagc	agtttctcta	tagaggggaa	gcaaaaggct	171600
ggagctgcgg	tggtagacgg	gaaaaaggta	atttgggcaa	gcgctttgcc	tgaaggaaaca	171660
tcggcacaaa	aggctgaact	tatagcgctt	atacaagccc	tccgagaggc	taaaggtaag	171720
atcgtaata	tctacactga	cagccgatat	gcttttgcta	ccgcacacat	ccatggggcc	171780
atctacaggc	agcgagggct	attgacctcg	gctggtaaaag	acattaaaaa	caaagaaaaa	171840
ttcttgcc	tgtagaagc	catacatgca	ctgaaaaagg	tagccatcat	ccactgcccc	171900
ggccacccaa	aaaggagaaa	acttggtggc	caagggcaac	cgaatggcag	acttagtggc	171960
aaaacaagtt	gctcaagggg	ccatgatctt	aactgaaaaa	ggtgatccgc	ccaaaagccc	172020
tgaggatggg	aggtataaca	taaaagagct	atggtagacc	agtgatcccc	tcccatactt	172080
tttttgaaag	aaaaatagaa	ttaactcccg	aagaaggaaat	aaaattttgt	aaaggactac	172140
accaattcac	ccacctggga	gttgaaaaaa	tgatgagact	aattaaaaat	tcccgatacc	172200
aagtccccaa	cctgaagtca	gtggctcaaa	agattataga	ctcctgcaaa	ccatgtgcat	172260
tcactaatgc	aactaaagcc	tacagagaac	ctggaaaagag	acaacgggga	gaccatcctg	172320
gagtgtattg	ggaggtagac	tttactgaag	ttaaacctga	aatgtatggt	aacaagtatc	172380
tgtagtatt	tgtagacacc	ttttcaggat	gggttgaggc	atttcccact	aaaacagaga	172440
ctgcccagat	tgtggccaag	aagatccttg	aagaaatcct	gccaagattt	gaaatcccta	172500
aggtaatcgg	gtccgacaat	ggaccagcct	ttgttgccca	ggtaagttag	ggcttggcca	172560
ctcagttggg	catcgattgg	aaattacact	gtgcttaccg	ccctcaaagc	tcaggacagg	172620
tagaagaagt	aaataggacc	ttaaaagaga	ccttgactaa	attagccatt	gagaccggca	172680
gaaaagaactg	ggtggctctc	cttcctcttg	cgctcaaaaa	ccccgggtcg	tttcgggctc	172740
actccttttg	aagttctgta	tggaggacct	cccccttaa	tggaagctgg	tggaacatta	172800
gtttccgact	ctgacctgt	cttaccctcc	tctttgctta	ttcatttaaa	ggccctaaaa	172860
gtgattagga	cccagatttg	ggaccaactg	aaagcagcct	ataccccagg	gaccaccgca	172920
gtacccacag	ggttccgagt	tggagacaaa	gtcttggtca	gacggcatcg	aaccggtagc	172980
cttgagccac	ggtggaaggg	accctatttg	gtgttactga	caaccctac	tgcggtaaaa	173040
gttgacggaa	tcgcctcctg	gatccacgce	tcccactgca	agagggccgc	cagtcgaagt	173100
gaagaaaaac	acgacgacaa	ttggacagtg	gcagtcactg	acaatcctct	taagcttcgt	173160
ctgcgccgca	ggcgccactc	tagacctagg	gaaccttaac	cctcatgctc	caattcaaca	173220
gtcctgggag	gtgcttaatg	aaaaggaaaa	cattgtatgg	gcaaccactg	cagtccatcc	173280
cctctggatt	tgggtggcctg	atctcacgcc	tgacatctgt	aagtttagcg	caggatcccc	173340
caattgggac	ctctcagatc	atactgatct	tagcaaccca	ccccctgagg	agcgggtgtg	173400
cccaaattgg	atagggagca	catatgggtg	ttcggggcag	ttctaccgag	ctaacttag	173460
agctgcacat	ttttatgttt	gccctggctc	gggtcagagc	aaaaggcttc	aacaaaaatg	173520
cgggggggca	tcagattact	tttgtggtaa	atggacatgt	gaaacgacag	gagatgctta	173580
ctggaagccc	tctctaaat	gggacctaat	cacggtaaaa	cgaggtagtg	gctatgataa	173640
gtcaaacgaa	ggagaaagaa	accctataaa	atatcaagag	agtgggtgcg	cttttaaaaa	173700
cagagcaccc	tcaggaccat	gcaaagataa	atactgtaac	cccctacgta	taaggttcac	173760
cgagaacgga	aaacaacacc	gtctaagttg	gcttaaggga	aatagggtgg	gttggcgagt	173820
atacattcca	ctaagagatc	ctgggttcac	tttcacgac	agattgacag	tgagagaccc	173880
ggcagtgaca	ctcgtagggc	ccaacaagg	ccttataaaa	caggggcccc	ccagtcgtac	173940
tggctcccc	aaaggtcccc	actgtaccag	ctccaccaac	tccacagccc	aacacagtgg	174000
taccctccct	aggaactaat	actctcctca	taaagcctac	cttggcttcc	ccaccgcccc	174060
taggaacaga	ggaccgtctg	gtcagtcctg	tccaaggagc	tttttttagtt	ctaaatagaa	174120
ctaaccctaa	tatgactcaa	tcatgctggt	tatgctatgc	ctctagcccc	ccttattata	174180
aaggaatagc	tcagatcagg	acttataata	ctactctaga	tcattctcaa	tgcttttggg	174240
gaaaaaacag	aaagttgact	ctagcagcag	tttcagggaag	agggctttgt	ctgggcccgg	174300
tacctcagga	taaagggcac	ctctgtaatc	agaccagaaa	catccagtct	agcaaaagcg	174360
gtcagtatct	ggtgcctccc	ctagacacag	tgtgggcttg	caataccggt	ctcactcctt	174420
gtgtgtctat	gtctgttttt	aatagttcca	aagatttctg	cattttggtt	cagcttattc	174480
ccagactctt	gtatcatgat	aatagttcct	ttttagataa	atttgaacat	cgggtccgct	174540
gaaaaagaga	accggttacc	tttaactttg	cagttctatt	aggattggga	gtagcagctg	174600
gagtaggtac	aggaaccgct	gccttaatta	agaccccccc	aatactatga	agaactacgt	174660
gcagttatgg	atattgatct	tagaactata	gaacagtcta	taaccaaatt	agaagaatct	174720
tttaactccc	tgtccgaagt	ggtgctgcaa	aatagaaggg	aattagactt	attattcctt	174780
aaaaaaagag	gactctgtgc	tgcccttaaaa	gaagaatggt	gtttttatgt	tgaccattca	174840
ggagtaatca	aagattctat	ggctaaactt	agagaacgcc	tagatatacg	taaaagagaa	174900
agaaaagagcc	aacaaagatg	gtttgaaagc	tggtttaata	agtccccctg	gtccaccact	174960
ctcctctcca	ctatagcagg	acctttaatt	acacttatgc	ttttgcttac	ttttggggccc	175020
tgcatacctta	ataagttagt	agctttttatt	agaaaaagga	ttaacgcagt	ccagggttatg	175080

gtactaaggc	aacaatatcg	gtccttcag	gaggttgaaa	actcgctcta	agattagagc	175140
tatctcctaa	agaagtgagg	gaatgaagaa	taaaaatfff	tactgaactc	ttcttcaccc	175200
cagaacccga	ccccctccat	ctagagattg	ttcccggaa	actcctgaac	tcttcacccc	175260
agaatgcatt	cctgaactcc	tcaccctaga	gttcgaaccc	tcccaactaa	aaactgttcc	175320
tagaagcaaa	ttgagataag	ggcctccctaa	aacaaccgca	aaatgaaccg	ggtacattgc	175380
caaataatag	gacatgaccc	cttagttacg	tagattccct	tggcagaacc	ccttgtcccc	175440
tgacagaacc	ccctagtgat	gtaaacctgt	actttccctg	cccagctctc	cccccttgag	175500
ttttactata	taagcctgta	aaaaatfttg	ctggctcgctg	attctcctct	acaccactag	175560
gtgcatgagt	ttcgaccccca	gagctctggt	ctatgttcca	tgtgctttct	tgctgttggt	175620
ctattaaatc	ttgccttcta	cattttgagt	acggctcag	tgtcttcttg	ggtccgcggc	175680
tgtcccgggg	cttgagtgtc	tgagtggagg	tcctcctctg	ggggtctttc	aaaactactt	175740
cagaggaaaa	atgtattctg	cctcatgggt	tccgggggtt	tccctcagca	aattcaggga	175800
agacaagatg	gaacagctca	acctgctggc	aggagggtgt	gggaaaggac	aagtgttcat	175860
tgtgtgggtg	acaggaaaca	gagagctgcc	tacagtctta	caggcctacc	accactgacc	175920
tacctctgtc	cgtcaggccc	tacatcttaa	aggatctaca	gtttattaaa	agaacactac	175980
cagataggaa	ccaagtatca	aaccaccagt	ttgtagggga	taaaaataca	aggaacacat	176040
ctcaatagga	gtgtgttcca	ggatgtggac	aaggagaaca	cagttgttta	aaagcttaac	176100
gctggccagg	agagctgcac	acctttaatt	ccatcactcg	taagagggaa	gcaggttcat	176160
ctctgtgagt	tcaaggcaag	cctgggctat	acaattctag	attagccaga	gctacatcgt	176220
aggagcctgt	ttcaaaaaca	acaaaaccaa	accataaaaa	agcatttctg	aggctttggg	176280
tttaatcccc	atgacctcaa	atagccaaac	agctctcctc	agtccaaacc	aaactgcaaa	176340
attggagcta	gtgagatggc	tcaacatatg	aaagtccttc	ccaaaaatat	tgacaactgt	176400
agcttatctc	tggggacaca	cataatggga	gaggaccaat	ttctacaagt	taccctctga	176460
cctccacaca	tatgcctccc	acaaataaga	aaatatatat	aataaaaaga	aagaagtcta	176520
cagctgcaca	tggtcatgca	tgctataatt	ccagcactcc	agaggctgag	agaggaggat	176580
tattagtttg	agatcgcata	gcaagcagta	ggctagacag	ggctacatag	tgtaaacctg	176640
ccttaaaaca	caaaaatcaa	ttaagcaaca	ataacagtaa	caaccacaac	aaaaacccaa	176700
aagagtactt	tgtagtaagg	acaataccaa	aatgttccct	ttaaggacag	ttctgggaatc	176760
agcaatagcc	ttccgagtg	tcagggatgt	ataaataactt	agaaaacttc	ccctggagaa	176820
atgagcacca	gggtacactg	ctctcagagc	tgcccagaaa	gttggtttatc	ctggattcat	176880
ttcagccttc	ctaactgtc	aggcattcag	aggtcacttc	tgtagtagcc	aatgtctaaa	176940
aaggctaaac	tactgtcag	catggctgtg	ctacttgcca	ttatcatttt	gtgactgggt	177000
ttgtagttat	gcagaattca	agagttatag	catcatgaaa	gtttccacca	agttcctgat	177060
ccagtcacct	cttaaagggt	ggatgcacca	agtgcctttg	gggtgataaa	ttatattcaa	177120
ataatgggat	tccaccctaa	tccccaaaaga	cttctggcca	tctcataatg	taaaatgctg	177180
agccatcgca	ccagcccatg	gccttgaact	cttgatggtc	ctgtctcagc	ctgtgttttg	177240
attataaatc	tgttggtgag	gtattccttt	gctgataata	caagcaaatt	cttcaagctt	177300
ccatcctaga	tgtgaagacca	gcagctctcc	aggagctctc	aatgcagact	ggccagctg	177360
ggacattgag	cctcatggac	tcagccgcta	ctagattcgc	aacctattca	gacaagccac	177420
tgttgagcta	cccagacaat	actatgtaag	ccaatcccat	tttaatacac	atattcatct	177480
gggtgtgtgg	cacacacctc	tactcccagc	acgcaagagg	cagaggcagg	cagatctctg	177540
atttcgaggc	ctggtctata	gagtgaattc	caggccagcc	agggctacac	agagaaaacc	177600
tgtttcaaca	aaaccaaaac	cgtaaattca	ttctatcagc	tctatttcct	tagagaattc	177660
taatacatgt	gggtaccagg	ggttgaactc	aaagtcttca	tgtttacgta	gcaagatttc	177720
ttctgctagc	ctagtgaagc	tgaggcaggt	acagccggtt	ctttactgct	ccttgcaaat	177780
ggtcctcctg	agctttcctt	tgagagccta	caaagaactc	tttttttctt	taggtctcca	177840
ggttttggte	ttaagagggt	ctggacttgg	atctgtagct	gtcatatcac	agacattcaa	177900
catctggcaa	atgtcttgac	aaaggagatc	acttgtgttt	gctgcagtg	cccttctggc	177960
tgtgagattt	tgctcctcac	cactgcagga	ctgcagatct	attctgcctt	tttagttgac	178020
ttttcattcc	tgagaactgg	ggaaaactga	ctttgtattt	gggctttgaa	tttgtccatt	178080
tgtcaatcca	tcacaccaga	cctaaccaac	tgccaagagt	tctgctgact	tttgttttct	178140
ctagggtggt	cactttgctg	ggctcatcct	catccttggc	ctgcagttta	tccccaggaa	178200
agaaaatggc	taacgactgc	taagaagcag	tctttccttc	cagaaatfff	agtctatcta	178260
gaccttgctg	cagtctgaag	tctttaaaaat	gtgtttgtta	tggtagaata	ttttgagttg	178320
ccttaggagt	attgcttgct	gtcacctatc	atattctatc	aggaagcaga	cgtcccattt	178380
accaaattgt	aagaaatatg	gcatcaatac	ccactgcaaa	aagtgtaaat	aaataataaa	178440
aaaatagatt	tattacagag	tgcaagggaa	acagaaaaaa	tcagccagtt	tcagaattgt	178500
aactggacaa	atgttggtac	agttcatgaa	gaggttctac	aaaatggctg	ggggtgggaa	178560
cataatgagt	tagtttgctt	ttttttttct	ctttccttcc	ctttcctttc	cttacaagggt	178620
ctcatgtagt	ctatggtctc	aaactcacca	ctgtaaatca	ccttgaactt	ctgatccttc	178680
tgacagctgg	caatgtaagc	atgtgccacc	aggcctggct	cacacatttg	gtttttcaat	178740
acagaatagc	tctgtgatga	ttaacttcaa	tcatcaactt	gacataacca	agaatcgtct	178800
gaggaagagt	ctcagtgact	gggtgggcta	agggcatgct	cataagggat	tatcctgatt	178860
gttaattgac	atggaaagat	caagtcattt	gtgagcagca	acacgccctg	aacagaagtc	178920

ttctgaagta	taagaggaga	aagcttgatg	agagcaagca	ggcaagcaag	ccaggatcca	178980
cgtgtttatt	ctctgtctgt	tcttgaccgt	agatgtgatg	gctgtcttgg	cttcctggga	179040
aacatgaact	gcaccctgga	attgcaaggc	aaacaaacct	tttcctcttc	caagttgctt	179100
tatgctaaga	tattttatcg	cagcaataga	aatgaaacct	agaacaggcc	cataactgcc	179160
agctttggaa	ctgaacctaa	ggctgttata	attcacatgg	atagggacca	ctggaagtga	179220
atctgatttt	gatggtaaaa	tcatgtgttt	gtttctggat	atgatagatt	tatcaatttg	179280
agactcagaa	aagaagttag	gacttgaatt	ccgtttttaga	gacattccag	agaaaactga	179340
tgtcattgtt	ctgaatgtaa	gtgcctcagc	tgaaaataca	aagagtacag	ggaagaaagc	179400
ccaggctaga	atctgaagga	actcctctat	tttttgtttg	cttgtttgtt	tgggttggtt	179460
tttgagacag	ggtttctctg	tgtagccctg	actgtcctgg	aactcacttt	gtagaccagg	179520
ctggcctcga	actaagaaat	ctgcctgcct	ctgcttccca	agtgtctggga	ttaaaggcgt	179580
gtcctcaccac	accagctag	gaactcgtct	attacacatt	aacacccctc	tttaattaac	179640
tgttcctgcc	aatgtaccaa	atagtcaatt	gattcctgtt	tatttaccac	atgtttctgt	179700
tagtaaacca	gaataactta	tctagccaaa	gtctgcctat	tagccatatt	ttcatcagtt	179760
cccaaccatt	tttggaattc	tgtgagggga	atccacagat	gctgtagacc	gctttagaca	179820
tttttcagct	tttttcaagt	tgcaggtcat	gattcagtg	gtcatgaaat	taatttagtg	179880
ggttctgatt	agcatttcaa	aatgaggcaa	gcagagggga	tattgtcaca	gcacagcaca	179940
tgcggtaaag	agccacacac	tcttgcttgg	aggcttagtc	agtttctggc	tctaaagcc	180000
ccaggtttgt	ttctctatcc	taggcctctc	tcttaaatcc	caaacatagt	tagacattac	180060
cattggggca	cgtgcaactc	aaacacggag	tgtgactcct	ttcccatct	gcggttccca	180120
gatttgga	tgtcaccctc	ctcccttctc	cctagggtea	gttttacctc	tcacactcca	180180
caacacaaca	cctctcatct	caagaattgc	cattagggct	ggtgagatgg	ctcagagggt	180240
aagagcaccg	actgctcttc	tgaaggttct	gagttcaaat	cccagcaacc	acatggtggc	180300
tcacaacct	ctgtaatggg	atctgattac	ctcttctgtg	gtgtctgaag	acagctacag	180360
tgtactcaca	tattttaaat	aaataaatct	aaaaaaaaga	aaaaaaaaga	ttgccattaa	180420
atgtacctca	gagtccaaat	gcttcttctc	ccctgacta	cactcacgct	ggcctgagtc	180480
cattttctta	ttgaggttac	tgttctctct	cttctaccct	ggctccttct	gctgcctatc	180540
cttgacacag	cagacaagca	gttctttaa	gcagggctca	ggaccagtga	gactgatcgg	180600
ctctggtggc	acttcctgcc	atgactgatg	atctaagggt	aagcctagaa	cccacgagg	180660
agaagcaaag	gacctaactc	ccaaagccgt	cctctgacca	ccatgtgtaa	actgcacatg	180720
tacatgatg	cacatggtag	acacacatac	acagaagtaa	aaagagattt	aaattgaaaa	180780
tcattaaaaa	gaaaaatcag	ggctcagcaa	actttccgtg	tagaaaaacta	gagtagttag	180840
gctttgaaag	ccaagaagtg	gatattaatt	atagttattc	attatagcag	agatttctaa	180900
aaccttttga	caaaactaaa	aaatataaca	gagtgtattt	tttttgtaat	gtaagtttac	180960
taatggcagc	agtgggatta	gtttcttttt	tagattattg	ttattatttt	tattaattat	181020
tagtgttttt	gtgtttattc	atattccaca	gcatgtgtgt	ggaattggat	ttctgcttcc	181080
acctttgtgt	gggtcctaga	gattgaactc	aagtcaccaa	gcttgacacag	taggttggtca	181140
ggcttacaca	tgtggtggc	aggcttgat	ctttggaagg	caagcatttt	acttctgtgt	181200
ccagctcact	ggccttcttt	gtttaaaaaa	agaaaaaaa	agtccttttt	tgtttaatta	181260
gggttcattg	ccagtgtctc	ttatcttaaa	atcaactgca	aactttttatc	tggtaaaaag	181320
ccatcccttag	ctgtggctct	aggagaaaaa	catacagttg	gatggcttta	tcctgcaggc	181380
ttagtttgat	catctctctt	tgaagatata	atcagctcac	atcacactca	agcctctgcc	181440
aacgagtttt	ctacttctgt	tcaacaaact	acccaagctg	agcagctcca	aacaacagcc	181500
agtttgatc	ctcacagtcc	ggtgggtcag	aagctctaag	gggcgtggct	acctcgctgc	181560
tattgcctga	ccctgtctcg	tgcattccaca	ttcacatcct	ttcctgggtga	gtgtggttct	181620
ttgactggtt	ttgttccaat	tttttagtata	tgtgtctgtg	aaacaatctt	tttgctctctg	181680
cctccagact	gcagggatta	atgttcttga	ctgccacaga	gcactaatat	ttactgaaca	181740
tgtgatcatg	tgggtgctcag	cactcttgca	cccaaggctc	ggggaacatg	gaggaagagg	181800
gggtggaag	attccaagaa	ccagaggaag	aagaaagtca	gaggtgagac	tgcattctct	181860
agaaatgtca	gggacatttc	tagacctctg	aagtctcaag	aacaaggcct	gaaagtctta	181920
tttatatagg	ttaacctgaa	aggggaaaaa	attcttacag	gggtccaacg	ttagacaaaag	181980
aactctaagc	aactaaggaa	tgttgggggg	ggggtagtct	tccccaggga	acactcctct	182040
acctttcaag	ccccacccaa	gctggttatc	caaaaacaa	tggtcagtc	tgaagccata	182100
tacgcacaag	taacatcata	tggatgggca	gattgcattt	aggaatacac	acatacacac	182160
acacaactta	aaaagagagg	ccatgaattt	aagagagagc	aaagcaaaag	gggaaggggt	182220
acatgggaag	gttggaggca	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182340
nagagagaga	gcctattatg	tcgttggttg	cttctaatac	ttagaaaacc	actctcttag	182400
gctgagtcag	aactagccta	gctgagacac	tgtccacca	ctgtcccaga	gcaaggccat	182460
cgctgtccca	gatttgctct	tgggggacct	tgaatgaaa	gtcaccagca	ggctccggaa	182520
gctgcctcat	gctaatacag	tgaggttggt	aaaatagccc	tgcagtggtt	cctgggcctg	182580
cagctggggc	agagccacta	aggggagctc	ggctcttttg	agcagagtta	acagtcacat	182640
gtgctttttt	tttttttttt	aatgttcctc	gctttaggct	cagtgtctgt	cgctacttct	182700
aatccttgca	ataggctgca	agacaggcaa	gaatatcatc	cctgttttgc	cctcaggcaa	182760

attctgaagt	ctggcaaagt	aaagatgtgg	gatttgaaca	cagacttggt	tggccaaaga	182820
attctcactc	tacttctgcc	tgtgccacct	tcctctcatg	cacggggagg	ggaggggagc	182880
ccacctccca	tgctcagggg	ctaggaagtg	gggagaagat	ggatgtcctc	aaagcagggt	182940
gagaatgaag	tagaagccag	cttcaaattc	aaactaagca	atgtttttatt	tccatttccc	183000
tgaacataaa	gttcagttac	atttggttta	aaaaaaaaaa	tccctacaca	actggttctt	183060
gagaaatgtc	aagtgtctaca	attcagtgga	tgtggatgaa	acaatcaaaa	tgttgaacac	183120
ccccaaacag	atacaaacct	tcatacaagt	ctcttccaaa	ggctgggtct	gaaaagagcg	183180
actcatgttc	cagcccagtt	ggctccttct	catgtgagct	ccgacttcca	aagactgctt	183240
gcaccaggag	gaaatataat	agatgtcctt	tttaaggggg	gtgggggtctg	tctgacaacc	183300
ttcccacagt	actgtggata	cagcccagtt	agtagagttc	tcgcctagca	agcgtgcggc	183360
cctgggcttg	agatctagca	ccctaaggca	tagtgggtaca	tgcccatgac	cacagcactt	183420
gggacgtaga	ggcagaagga	tcagttcaag	gtcagatgag	gggtggggag	gcatttcctt	183480
aatgccagca	tttgagaggc	agaaacagat	gaatttgtga	gttcaaggcc	agcctggtct	183540
acagactgag	ttccaagaca	gccaaggcta	cacagagaaa	ccctgtcttg	tcaggaaaaa	183600
agatagtggg	agagaattca	aggttatctt	ggactgcata	agactttgat	tccaaaataa	183660
acaaaaatgg	agcatgaatg	cttgcaactg	tggacaatat	tgggttcata	catattctgt	183720
tttgtcacct	acataccaat	tatacaaact	agattcagct	gggcccactg	gtgcatgttt	183780
gttcccagca	tctgggaggt	agagatgggc	agagctctat	gactttaagg	ctagctgggt	183840
ctacaaagta	agttctagga	cagccaacct	tacagagaga	tacactactt	ctaaatcaat	183900
caatcaatcc	atcagtcatt	catgggctgg	agagatagct	cagtgatcaa	tagtgccatt	183960
cttccagagg	acctgggttt	gattcccagc	acccacatgg	cagctcacag	atgtctgtaa	184020
ctccaatccc	aagggacatg	acaacttcta	ctgggtctctt	tggtcaacag	gcattgcacg	184080
agcatacaat	atatatatgg	gtaaaatgct	atatatatata	aaatcagatt	cacaaatcaa	184140
gtacagaaag	agatcaacta	taatgaacaa	accataacac	atactgtttt	aaaagctacc	184200
tttctgtctc	gacatctgac	ttcttgcgtg	accagtgccc	tccacaatga	gacaagcaga	184260
tggtgtagtc	cctgtgaccc	aggattaggg	aaccactgcc	cctgggaagc	catttcctagt	184320
aacaggatca	aatcccacag	tgctccactg	taccccgcca	gaggacgtga	agcctccctt	184380
ccccggctgt	ctgtgcagca	ccgcccagat	gcttgggtgc	actgagtgac	catcgagacc	184440
caactgagga	gtgcctcagt	gctcctcagg	gcagtgtgca	attgaaactt	gcattgtgatt	184500
tctggaattt	gtcattttgat	atttccagac	tccagctgac	cttggataac	agcacagagg	184560
gcagaactgc	ataggaaaag	gggcactact	gtgactgtta	tcccctgcat	gattatagaa	184620
gggtctgtgc	tttctcttac	aaatcattgc	ggcctctctt	cacttccctc	ctttgaaagtc	184680
gaaaaaaaata	gatgcattcct	cacagtacag	gatggcaggt	acgaggcggg	ttccgggact	184740
gaggcaggac	tcattcatgca	gcctctttcc	ctcaactaca	ccgcccggcc	tgagaggttc	184800
cctctgatca	aatcagtttc	aggcctggaa	agaaaaggcc	actcaggctg	gggatgtggt	184860
tccattagag	gcctgcatgc	acgaagctct	ggattttgatc	ttcagcacgg	gcataagccc	184920
agtaggttag	tattctgttta	gcatggtgtg	gaggtatacc	tatctctgca	caggagacc	184980
agaagtctag	agttatcctt	gcacttacag	taagttcaaa	gctagcttgg	gcaaatgaa	185040
gttttgtctt	aaacaaaacga	aaacagaggg	gccggggaga	tgcttcgctg	tgctgagtca	185100
tttgctgcca	agtttgatga	cctgagtttg	gtcccttgag	cttatggtgg	aaggagagaa	185160
atgacgcttg	aactccacgc	acatgccaca	gcccacgcat	gaatgtgtat	acacacacac	185220
actaagtgga	taaatgttaa	aaataaataa	atcaaaggaa	tggccactca	aaatctacca	185280
tcgttgggaa	gggaggggaa	aaggcaggcg	aggagatag	ataaccctga	tatgaacacg	185340
gaaagagcca	tggtgccacc	aaagtgcgcc	agtgtgccac	caaagtgcgc	catgtgtcca	185400
ccaaagctgc	ccagtgtgcc	accaaagctg	cccagacttg	attacagatt	tggccaggga	185460
cacaggaggc	cagcaggagc	agccagggtc	cacctcagag	gtggagccac	aaacctggaa	185520
atgaaacgtc	tttccctttc	ttcagaccac	agcagtgaca	gctgtcctgc	agagtctgga	185580
gggctggcag	ggctcatcca	ctctagtgtg	cctgtggcca	gaacaggcct	cagtacagg	185640
tgtctttcca	aggtcttagt	gtctaattaa	ggttagcagc	caaattggag	agagaagggt	185700
gctggacttt	actctgctgt	aaggactttg	ggcattgttc	cattccgtga	tcaaatacca	185760
ctggctctgc	caaccaccat	gtcagtggtg	cttcagaggt	agaagaactc	atcctttttt	185820
gagaggtttg	gtctggctct	tgtctaattg	aaaatgcctg	gggcaccagg	ttaatgtcaa	185880
ctcaaaggca	agtgtctggt	cagcatgtgt	ggaatcctaa	gttcaatacc	catcagagcc	185940
ccaagcccta	gaggacagac	atgcttttaa	aaaaagtcac	gctttaaaaa	aattctgttg	186000
agggggctgg	agaaatagct	cagcagttaa	gagcactagc	tgctctttca	gaggaccag	186060
attccgtttc	taacattttc	atggtggctc	aaagctgtct	ataattcaag	tccagaggg	186120
gaatctgttg	ccctctctgg	ttttctcagg	caccaagaac	acatgtggtg	caaacatata	186180
cgcaggcaaa	acactcatac	atatgaaaca	ttttttataa	acctgctcgg	atgtggtggc	186240
tcattgccagc	gatgctctca	gcactcagat	ggcagaggca	gggggatttt	gtgttgagcc	186300
cagcctgagc	tatagaatga	gatgctgtct	caaaaagaaa	aaaaaaacaa	aaaaaaacaa	186360
aaaacaaaaa	caaatctgtg	ggcttaatca	ttcctagcca	gaaggagctg	gctccagcaa	186420
caatggatcc	ctagagctct	gctttgcccc	ctggtctgga	gagcttgctc	tagaaaaggaa	186480
ttctccatag	accacatttc	tattttgggg	accactctgg	gcctgcacat	ggaaaaatgg	186540
agagttgagg	tttcatggtt	catttttttt	gttggttacag	taatgagagc	tttgaaatga	186600

tcacaaaagg	aaaataggaa	aatatgcctc	ctaaaaagag	ccgaggcaaa	ttattatagc	186660
aaccgaatcc	taaaagggaat	gttcaagtaa	aaaaaaaaaa	atatggctca	tgcgaggttg	186720
ttatggcaac	tgacatttaa	aagtaacagg	atttgggcaa	gtttctttct	ccacctctc	186780
tgcaaagtct	tgagggaact	tcattgctaaa	attatgcttt	aatttttaatt	agccaaatag	186840
gtcataaata	tagcttattc	ccaaatcctt	agaattttcta	ccctgcaagg	agctgaacta	186900
ctaattgagga	aatattttcca	caaaaaccca	cttaccatta	agaaccccca	tccgatattt	186960
ttctaatata	atttatcaat	ttaaacacat	tgcataatgt	gccactctgt	agcattccat	187020
taaaatgata	aatagcaata	gtggtgaggg	gtgggggggca	aaaaccagga	gattaaacat	187080
atccaaagca	gtgtagctat	atttaaatac	ctcaatccat	tgtagaggaa	aacacactgt	187140
tctcatcgcc	agatacagtc	tagactcaga	gcagcatatc	cttgactgta	aggggtattaa	187200
taggacagac	gaaggggggc	aataagaaat	gacaggaaac	ttcagaagaa	ataaaatttc	187260
tattaggctt	tgttataaga	ttacatcaaa	gcagttcata	tgttttaatc	tggggaggaa	187320
aaaaagcaac	tacttggggg	ttgcgcctgg	gggctgcctc	tgtgtactga	accagacagt	187380
ttgcataatg	aacaattttc	attcaatcag	gatctcagca	gagatagctc	ctactcaaag	187440
gaacccggca	caggctcata	gtttttatct	cccagctcca	cctgctggag	aaaccttgta	187500
ttgcagggag	agaaagcagt	cgggaggcat	tgtcctagt	gctgtgtacc	taaagttaca	187560
gacctgactt	taaacagttt	ctctctggag	gttgaaaggg	gctctgtaag	ataccagagt	187620
ggattgctct	caaagactct	cggactcctg	ttacaggcaa	gtaaggctct	agcagatggg	187680
agcatggatc	tccggccctt	ctcactgctt	tcttgaatca	gggatttaga	aattgctatt	187740
tgcataccag	gaggactgaa	gtttggctcc	cggtgaccag	aggacaaggt	cattgttttaa	187800
aaccacccaa	actcatttcc	gacttggttg	gtcaaatttt	caagtttccc	agcagctctaa	187860
ggattcataa	aataaggcag	aggcagagaa	acggagggtg	tgtgtgtgtg	tgtgtgtgtg	187920
tacccaaaaa	ggaactgcat	tttcatgcac	aatagaaaac	ttaaagactg	aaccaatcat	187980
tttgaaaaac	tggcacagct	gacattggct	agaggaaagga	acggccaggg	cgagccagct	188040
gcaccaagac	ccagggtctga	ggcctaattc	gcctttatcc	gagggtttag	tgaggctccc	188100
gccgctcacc	aatcccggct	ggagccgcag	aagagctctc	ttcacttggc	tcagtcccag	188160
cacagtcgca	ctatgctctc	ccgtggggag	gccgctccgg	gagggggagc	gacatcaagc	188220
tttgtgaaac	tgttttcgaa	aacctgggat	gatcatttaa	atgtttaaaa	tatgcacatg	188280
gtaattcaaa	actaattacc	ctgagcacat	ttgaaacatt	tatgccatca	tcttggatcc	188340
tgectactga	ttgtgcgtg	cagctcactc	tggtgtttct	ataaactgct	tcagcgattt	188400
taacttccag	gctaaatcag	gcagccacag	gcgctgcctc	cagccctggg	ttgggtggaga	188460
gacccccatc	cctgacttcc	aggcgaggag	gcgggccggt	tctccagaga	gccgtttgtc	188520
aggggtcttg	agttctggct	gccgaattat	tgctcttatc	cgtgttcata	attctcatct	188580
gcattatttta	atttaggcta	gaatgacctc	tttccctccc	gagtcttcc	ccctcattcc	188640
catttccctc	tcttcaattc	gtggccccca	ttttctgatt	ggtccaaata	tatagacaaa	188700
tatccttgat	cgtcccaccc	cacttggcta	catcttcatc	tgggagccaa	tgtgggtgagt	188760
tttctgggtt	tgcaagggtg	tcagggtccac	cagtcctcct	aagggtgtgtg	agagaggtag	188820
accaacatga	gcggcgca	gccgccatca	ctgagagagc	acgtgccctg	cagctcaggc	188880
acaggcatgc	acacaccggc	agacatgtgc	acatgcgctt	tccccagcaa	accctgcttg	188940
cagagtaatt	aggcctaggc	agttcctgaa	gcaaattcat	ttcccccttt	tccagaataa	189000
aatgagttct	cttcccttgg	gggtgctaaa	ccagcatgcc	agtggctaga	agcctgagat	189060
gggtgatgtg	gctgaaacca	tttctgcagc	caagcctgtg	ggcagaagct	aaccttgggc	189120
tggggagctg	cagtcggaag	aggcacaatt	ctgggagcaa	gaaatgagca	ctgggttata	189180
ggtagactcc	cagaaaataga	cagatgaggg	ctgcctcctt	attaagcgctt	tgaagatgcc	189240
catggcgggt	ttttagacat	ttaggaatat	aaaagtaggt	tggattccca	cagtcagctg	189300
aagtttgaca	gagtgatatt	accgggttta	actagagcca	ttaagagact	cttcattatc	189360
ccacaccacc	gccacccaag	ttatcacatg	agccataatg	caagagaatt	ttcattccat	189420
caacaagaga	gggagccggg	ctatctttgt	ccaaaggaaa	tgagcagccc	agcgtgaagc	189480
ttgtgaggaa	ttgagtgtac	aacactccaa	taacatcccc	tgaggatttg	cctctgcgat	189540
ttagtgggtg	aagcaggggt	aactgcgctc	gagcagctctg	cctgtgtacc	tggcttgcaa	189600
gaacaccagc	tcgagggaaca	ccaaaaaggc	cgattaatga	caaaggacac	tcatagaggc	189660
ccgaattcca	cagggtctaa	gtattaagcc	ccaaagaaat	caaggtctag	gccattctcc	189720
tggcgctcag	caatctcatt	tattatttct	ctacaaagat	ccaacactca	atttcccagg	189780
tatccccctg	atctgactca	cattctcctg	ctcagtaagc	catcctgggt	tgaacggggc	189840
ctccccctct	cctgcctatg	catgctttgc	gtcttcacaa	cgacagctgg	taatttgcaa	189900
gacccccctc	ctcaggactct	ctcaccccc	atacttggaa	ctactccttg	gaactcttg	189960
tttatcaagt	gttctgttgg	tgagccttct	cttgcattaa	agctgtgaga	aggaaccaca	190020
gtttctatct	cctttacatt	tcttgtagcg	tctcacatgg	gagacacca	ggttagatat	190080
actgagggtc	ctggtagttt	tagagttgga	gttagatgac	ccagcaacat	gccttcccc	190140
accacgcacc	aagcaaaaaat	tgcacccacc	cttccctcag	atgttcctgg	catcttataa	190200
ctcgcccaaa	gccagattta	ttgctcctgc	tgtaaagtgt	atcttctcta	agcctcactt	190260
aaaagctacc	acttggcaga	agatcaagtc	agaagtgcag	gctagcaggt	gacggtaggg	190320
acagggcggg	atggggcggg	tagggtaggg	cgaataattg	aagctccaag	agttaccagc	190380
tcaatatatta	acctaactgg	taatttgctg	tgacaattac	gccatgaagg	gaacgctgag	190440

actatgcaag	aatgttgctc	tctaattaag	agggctctgc	atttcctagt	cacccgcact	190500
ttaataacac	acagaatgag	ccttgggctcc	gggagctaaa	ggttccatta	ggagcacggg	190560
cagcatatgg	ctgtgcacat	aggccgtgag	tgatgcagcc	cagttaagcc	cgctaacacc	190620
ttcaattcgt	cctcagatag	agcccagaga	gcgcggctca	ggccctcacg	ccacgagccc	190680
catttgactg	acaggtatct	tcccggaaag	cctgcgcgtg	cctacactgc	aaatggacct	190740
gcttcccaca	gcccggcttt	caaccaggaa	ggcttggcgt	gggtctgate	cttcaagagt	190800
aactttaata	aggattttct	cacagaaaga	aaagtccatg	ggaacaaatc	ctcctcttaa	190860
gagcgtgaga	caggaatggg	gacacaagcc	aacaccccaa	ttgctaggct	aactctgata	190920
tgagacaaaa	gaatattaat	atcttggcta	tgaaggagga	tggtgccatc	ttctgaattg	190980
atgggagttt	tgaggcatgg	ctaagctggg	caaaccattt	tctttttttt	ctcttcttaa	191040
ttagtggttc	atttatggag	ggcttgcctg	ccggagagcc	catcagaaga	gagctcgctt	191100
tatggagatg	tagcttataa	aactactcag	attttaaaca	aacagtgcag	gagggccagag	191160
gtagaagtgg	tgggggtggg	gtggggcaag	agaacaattg	catctgcaga	aggctagccc	191220
tgaccccaa	gcctatgttt	agggttgatc	agcttcccga	ggcaagccca	gaagcctcta	191280
aaatttttagg	ccaatagaaa	tgacctctgc	accacggctg	actgaagcta	taaataagcc	191340
tcgagttgag	cagtgggtgc	aacggagaga	gcagaggaaa	gtccaatcag	agcttcattt	191400
tttttttttta	aagtccactt	gcttgggact	cacctgaagg	cagggcattg	agtagagcct	191460
tggctccctg	cagcgagagg	ctccagtttt	cccagagacc	agcccatcgg	ttggttacct	191520
aaccaccgga	agggaaactgc	acagcacaca	agttaaatat	aggctgggtt	atctgcattt	191580
tacaagctct	gagcaagcta	tctgaagaag	ctgtcatttt	taatgacggc	acaaacttcc	191640
aattaccgac	tgggtaatcc	actagggagc	aggtagtttt	ggaagaacag	ttcaccatta	191700
ttaaaagttt	acacaatcac	ttttgagttg	actataagta	tttcacacga	ggcaggtggg	191760
attagggact	ttttgggtgg	tttactcgag	gctgcaacca	acaatgagtg	ttttctcaag	191820
aattatacat	tgagattttgt	caactgctgg	ggagtagtgg	agggctcctg	taatgcagaa	191880
aggtttatgaa	tgaggcaggt	aaggttgggt	gcttccaagt	ctcaaatata	ctcctaaggc	191940
cagctccaag	tcataagctc	aaacaagtct	tcaaggggcc	tggagagtta	agacaaaata	192000
ggatcactta	ggctacccac	ggacaagcac	ttctcataca	aggaccggct	acctccaaca	192060
ccatcttccc	aacatggctt	ctatgttgct	tcaacaacca	gggcaggggtg	aattaggggt	192120
gggtctctcc	aatgtggact	caaatcatga	ctacagcntg	gggttttttt	tttttttttt	192180
tttttttttt	tnnttgggtt	ttcgagacag	ggtttctcca	tatagccctg	gctgtcctgg	192240
aactcacttt	gtgaccagag	ctggcctcgc	actcagaaac	ccgcctgcct	ctgcctctgc	192300
ctcctgagtg	ctggaattaa	aggtgtatgc	taccacgccc	ggcagagtc	gtcttgataa	192360
tgaagtccc	agtgaacctg	atgtcaactg	aagttggatt	ttactgtgat	gactactgag	192420
tccggctcag	aattttgggg	ggacaaggta	ccttgattta	actgggcact	acacgactgt	192480
aacccccaca	ttgggagagg	cagaggcaga	ggcagaggca	gaaagtgggt	tggaggctag	192540
gcaaggctac	acagcaagaa	gctgtctcaa	aaccaaagac	atctttcttg	atccaaatcc	192600
tgctggaggg	tgtgaggcct	tggggggccag	aacaaggctg	tcaaggaaga	ccactgactc	192660
tgtcctttgc	tccattactt	aatcagaatc	gccatcacag	atatagctag	gagattttta	192720
gccttggtgg	ctgcaatctg	catttaagag	ctaagtggga	taaactcagg	ggtggggcca	192780
atgcctccct	ccccaccctc	cctgcacccc	tccatttacc	tgtttccagg	gatctgctta	192840
atttacctgc	cagccttttg	tgggacacag	gcttagtggc	ttagcgctgc	tcggggcacc	192900
agagaccctc	acagaagcac	ctgaatgtac	tttcagcgct	gcagagcacg	cacggctcag	192960
gcccatacaga	agaacccagg	cttatgctaa	ggagccagaa	agtagaagca	gctggcaaga	193020
gtgattcagc	ccataaatt	tacacatccg	tacagccaaa	cccacttgaa	gtgatccaga	193080
gccactttta	ttgaaataga	aaagatgcct	attctggagt	gctaagtggg	acaggagggt	193140
gggtatataa	gagataatcc	catgttgtct	ttgatgtggg	gctagggaga	taaccaggga	193200
cctcacgcct	gcctgcaagg	tagccaccaa	gccacacca	caacctctat	ttatacacac	193260
actaagtgtg	gaggtatgga	taaaaaaaaa	tgtcccaaga	cctcacgaat	ctgcaaacat	193320
ggtgcctggg	tgggtggcacc	gtttggggag	gcagtggacc	atlttggtctt	gcaggaggaa	193380
gttatgtcac	tgggtatggg	ctttgagagt	ttgtagcttt	gctccccttc	cagttaaactc	193440
tgtctcgtga	aggttcctgc	caccatgttt	cctctgccat	tatggacacc	tggtcctcta	193500
gaactgtaag	ccacttactc	tcaggtctct	ttcagtcctg	gagtcttatc	atagcaatga	193560
aaagtaactt	gtgtggcagc	cagctaagca	agggctgtgg	ccgactgctt	gggattatgg	193620
ttgtgtctgt	ctgtctgtct	gtcattccat	ttatatagtc	ctgagaattg	aaccacttta	193680
ccactgacat	gtctcagtc	tcttgggtatc	atatattcac	ttaagacaag	atctcattaa	193740
gtcattcaga	ctggttttga	gcttgcaatc	ctcctgcctc	tgccctcaagg	cgataggatc	193800
cctagggtac	tcgaccagac	tgggagttagc	aggttctgtt	ctcttagctt	tctacagtga	193860
ttgtggatta	tttgtgtata	aagatctgat	ggcccgaccg	actcccttcc	ctttaagtga	193920
acatcaacag	tatttagcat	caacttaata	aactcatttg	gtaaagccat	ctccccacct	193980
cttgaaacaaa	tgaaaatcaa	acagcagtac	ctgttctcct	agagcagcgg	ctctcagcct	194040
tccggccttt	taatacagtt	cctcgtgttg	cggtagcccc	ccccccccc	agccgtagaa	194100
ttatttcatt	gcttaaccag	agttaactgg	aagggttaat	aataaaacca	gtctgggaga	194160
ctaaggttac	ccaaccacgc	taggaaggga	aggaaagggc	cactcgaca	aactgtctt	194220
tgagatgaag	aacaatcaac	ataacaggga	cagagcagtc	cttghtaaca	gtgcaaaggga	194280

gagagagagg	ctgagtttct	acttctataa	ataaacccctt	ggcaggcgga	tcactaaagg	194340
aacacaagtc	aatataaacc	tttagacatg	gggctgccaa	acttcacttt	tcgacagtat	194400
attaattatg	tagtcaatag	ccatggggtt	cattagcgta	ttaaataacca	cgatcaatat	194460
tattttatact	tttcgaagac	aagccactca	gggaaaaaat	ggtgggggga	ggaggaggaa	194520
caatttgacc	ctgtagttca	aaaaaagtca	gaacagcaca	ctagagatta	gcaaggggtt	194580
aatggaaggc	ataaaacact	ggaaatatgg	acagaaatca	gatccctgcc	ttcatttttc	194640
tgctttttac	aaagagactg	gaggggaattc	agaaactatt	taaaataaag	gcaaaatgat	194700
tagagccctc	ccctcccctc	agctgcttaa	cactgggggt	gtgggtggacg	caaaataagc	194760
attgagctct	aagtgataga	tgagaatcag	aacagggaaca	gtgtttttga	ggcaaaatat	194820
gtccaagaga	attcaaagaa	ctgtgggcca	gaatctactt	aggcagtcct	ctgggacccg	194880
aatccctcac	aggcgttaac	agtggaaacca	atttccaagg	cagccctgct	ggtgatctga	194940
tttttgagta	gggaaatctg	ttaaactctg	tcccacgagg	gagcccagct	ctttcactcc	195000
ccacgggttt	ctacatgcag	ctgtgctaga	tctgctgaag	tggccgggtga	ggaggtgtgg	195060
ggattgggtc	agcgacctca	gaggacattc	ttgttacta	gccctcgtgc	actggggcga	195120
tgaccgaatg	ctgtgagcag	gagatatcaa	aggccggcta	ctggactgaa	aactagatca	195180
ccatctctaa	cctgcaattt	gtcaatctca	gacagcaatg	aagactgtga	ttttctagtc	195240
aacgctttgt	aagcaaggtc	agatagaggc	tccataaaaa	ttgttcaggg	ttcaggcgga	195300
gaatcaagtg	taactcaatc	cctatctcct	gagattaggg	aaggggaagg	agcctgtgtc	195360
tactaaacca	tgtagcctca	agcaaagcct	gtctgttctc	agcaagggtga	ggcaccacc	195420
aaagatgcca	acagctaagg	gccagggatg	tagtgcaggg	tgctgtgata	tcaacagctg	195480
ggagacagaa	acaggaggat	caggacttca	aggtagtttg	ggctataaaa	tataagcttg	195540
aagctaccca	cttgaagact	gtccccaaca	aaacaaacaa	gctgggtatg	gtgggtcgatg	195600
cttgtatttc	tagtgtatga	gacgaaggaa	gaaagctcaa	gcccagagacc	tgcttgggtt	195660
acatagggaa	gatgggtgct	caaaaacagg	acagccgagg	agcagacaga	cagggcagac	195720
aggggtgcac	gatctaaatc	cacataactg	gattttaaagt	aatatctggg	agactggtct	195780
gtgagggccg	ttccagagat	ttaataaaga	cccaccctga	ctgagtatgg	gcaacaccca	195840
tgggtggccc	aggggtccag	actgaataaa	gggaaaatgg	gaggaagtgc	agctggtagt	195900
gtttccaagt	attaggacca	cagcctggcc	cctggcatgt	gctggccagc	tagtctagct	195960
ccagtatcaa	gcttcaggcc	agcggcaggg	cactggacag	ttcccacaca	cgacacacac	196020
acacacagag	cactagcatt	cacctcctg	tctcttcttg	acaacagata	aaatgtaact	196080
ggctgccaca	gtgagagtcc	cctaccttcc	tcaccgttaa	ggatggtaca	aactgtgagc	196140
gcagcagcag	catttctcca	gtaacttgct	ttccacagat	actgttatag	cactaagaaa	196200
agcaactgaa	acatgggggtg	ctgtgacccc	ttggcaccac	aaagccatgg	caagctgaag	196260
tgacacacac	acaggccagg	cctgaagatg	ctgggggact	gcaatgctgc	ctggattctg	196320
gcagagatgt	gcagcagatg	ccaagagggtg	ggctgcagca	accagagata	attaatatga	196380
ttaggaacac	actgagcagg	catgctcttg	ccgaatgaaa	agcctcgag	tgtaatgact	196440
gttttcttcc	tcgacacagg	tctccacgtt	tcagagtgtg	cttgggtgta	ggctgccg	196500
taaacatcaa	tccaaccccg	aggggcccaga	tcacggtgtg	tcctgggctc	aatcgctttt	196560
ccttttgtgt	tttcattcat	ttaaagatgc	attccagggt	tgcaaacatt	agtgagaatc	196620
atctccaggc	ctcagtcata	tctctgagtc	tgtaatgagt	taacatcttt	ccctagtga	196680
tatttattat	gaaggcta	taattgcttt	ccagttacaa	gaatccttta	cagtcaaaga	196740
aagtaggatc	cacaaagata	tactgtttat	tcaaacaaag	caaaggaaac	aaagcttctt	196800
tcttaaatc	tatttaacat	agctttaata	aaggtacaca	ggtccgcctg	gcaaccgaac	196860
ggtaactgat	gcaaacatgaa	gccatgctct	gtagcagcct	ggatgtccca	gtgccaccctc	196920
tgtctgcagg	ctttgtcgga	tttactaaga	tctctgttatc	ttcaaacagg	gattgtgtct	196980
caagtaactg	acccactat	gtggataatg	aagtaaatta	tgcaatttgg	gggtttgctt	197040
ttccccaagg	ggacagcaag	ccagtgtcta	tcagccgtcc	tcagaggaga	caattctgat	197100
taatatcaga	gtcatctgac	tcagtctatt	aaacctatca	aacctgaag	gaaggatatt	197160
cagatattaa	cgataggcct	ttgattaata	attctacctt	gttgccattc	taagcattaa	197220
caaccatgca	gtaactctgc	aaaacagacc	ctttgatctc	aggcagacgc	acctctgaa	197280
cacctgggtt	ctcccctact	cctctccccc	caggaggaac	tcaagacaaa	aaggtgccac	197340
cactggaaaa	gcacactcca	ggttacataa	tttgctcat	tatccagagt	gggtttaatg	197400
acttgtgaca	taattttctgt	ttgaagataa	caaaatttca	tgaaatccga	caaagccgga	197460
aggcaggagg	aggggactgc	tgccacacta	ccggtggctg	agaactggag	cggaagggttc	197520
acacacagcc	ctctgagctc	actgtctttg	cttatcagtg	agtcccaaga	ggggcccaga	197580
tgggttgcca	gcctccccta	gaggatcttc	attgtggagc	tgtcccatgg	ggcgggaagg	197640
aagccattct	attctgtttc	ttctctcttc	cgttctggcc	accagtggta	ctgtctccca	197700
tcacatgttc	ttcctgatgt	tcgcgatcag	ccgtctgcca	tagtctctga	agtccacggg	197760
cttcacgtcc	atcacagtgg	ccttaattcg	agattcatcc	tttagaaaag	agagaagctg	197820
tttgtgagtg	gcagagcctg	gcgtgcagcg	gaagagagaa	ctttctttgc	ttcagtggct	197880
tcaatgagtc	cagcaggaag	aaggaaagtt	tacaagtctc	agagagaaaag	tgctgtgact	197940
tcttgaggtt	gggccagatc	ctcttccaca	gaccttttcc	ccatcctagg	tgccctgtgc	198000
tcagacctag	cactctccc	gagaagcctc	tgtctttcta	tgggtgcagt	gggggcccag	198060
agcagacagg	taactcacc	taaagcatca	ctttcatcta	gaggagctct	gtggtagtag	198120

ggactgaggc	ttctgctcca	gctctgggca	aggttacttc	tctgctcttc	accattcctg	198180
tccatccag	gaagacagaa	aatccctaca	ctctcccttg	atctaccoga	ctttctgaca	198240
ccagcctacc	tatgttcatt	taatacaaca	actaaaatat	ctattcacag	gcactaagct	198300
ggtgataacg	cagaatgcac	aaactctgcg	gctgcagggg	agacggcaga	gttcctcctc	198360
cacttgctctc	cttgaactaa	acagtgtctt	tgaggcagaa	cagggtgaca	ccttagggaca	198420
cacaagtcta	gctggggggc	ttcatgcttc	catgtgctta	gtaattaatt	actacatgca	198480
ccgctgttta	caagtatggg	taggagcccg	actgcctggg	ttggcctctc	gcctctgcca	198540
ctccatggct	ttaggttcag	agtcattctc	tgcatgcctc	tgctgtctct	tccgttggtg	198600
aagcttgcaa	caacagctcc	aacacagaaa	gtgctgtgag	ggtcgacagt	ggatagatgg	198660
ctagatagat	ggggcaggac	ggactgtcca	gtaagcaggg	ttcatcatgg	ctatgcajct	198720
ctggacatca	ggattagttt	aaacacttgt	cagggtgggg	actttttacca	gcacgtgcta	198780
tttgtttaaat	attctgagtt	ttagaaccta	aactgtggga	aacaagagtc	cacacataac	198840
annnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198960
ctcctgctgg	ggttacaggc	ttgagccacc	atgacagctt	tagcaatagc	tttgtaaatac	199020
cacagtgtca	agctggatat	gatggcacat	gcttgcaata	ctaacctcca	aagattccct	199080
gaactcgagt	tggtgaaata	gtccaccagg	taaaagagct	tgctgcccac	gcctgaattct	199140
gattccctgg	tcacatgctt	taaggaaaga	acttgccgaa	gatgtcctct	gatgcccacg	199200
tgtaccgatg	catgcttgca	gccaccacca	caccacacaca	agtgcactgt	ctcacacagt	199260
gagaaacagca	agtgaacaaa	caaacaagcc	gggggggggg	ggattgtgac	cagaataatt	199320
gagggggggg	gtaaagctct	tggcaggtgg	ctggctcctg	gtaacactcc	ataagtgggg	199380
aagttccaca	tgtaagggtca	tgtgatcgag	tacatctggg	cctccaacag	tccttgngga	199440
agaaaacagat	gcagtctgtc	atatctaaac	cattgttgct	gtatctctgg	gtagtctttc	199500
ttttctcctt	cctttctttt	ttctctccct	ttctctttta	aaaaattatt	tatttattat	199560
tatatctaag	tacactgtag	ctgtcttcaa	acacaccaga	agagggtgtc	tgtcagatct	199620
cattatggat	ggttgtgagc	catcatgtgg	ttgctgggat	ttgaactcag	gaccttcaga	199680
agaacagtca	gtgctcttaa	ccgctgagcc	atctctccat	cccccaacct	ctttctcttt	199740
tgagttaggt	tttgtgtagc	cctgggtggc	gttaccttaa	ctacactggc	tttgaacttg	199800
caatgatact	ctgcctgatc	tgtcttaata	attttgagat	agggactcac	tacatagcct	199860
ttgctggtct	ggaactaaca	gagatctgcc	tgtttctgce	ttgcaaatac	tggaataaaa	199920
gttatgtacc	accacacctg	gagtttaagg	gttttttgtt	tgtttgtttt	tcgagacagg	199980
gtttctctgg	gtagtccctg	ctgtcctgga	actcgctctg	tagaccaggc	tggccttgaa	200040
ctcagaaatc	cgctgcctc	tgccctccaa	gtgctgggat	taaaggcgtg	tgccaccacg	200100
cccagtttaa	gggttttttt	ttgtttgttt	ttcctgaga	cagggtttct	ctgtgtagct	200160
ctggatgttc	tggaactcac	tctgtagagc	aggttagcct	tgaactcaga	aatctgactg	200220
cctctgcctc	ccaagtgtct	ggattaaagg	cgtgtgccat	caactgcccag	tgattttttt	200280
ttttttaatg	tgtgtatttg	tatgggtgtg	tgggtgcttg	tggaagccag	gtgtcagatc	200340
cccagagcgg	aagtgttctt	aaccgctgaa	ccatctctct	tccctcttcc	ctaaccttga	200400
ttttaaaggc	accaaactct	taggtaggag	actatacaca	cacacacaca	cacacacaca	200460
cacacccgta	cacacccgta	cacaccacat	gaccatgcct	gagcacacaa	gtggttttat	200520
tgctggctctg	gcctgtgtat	gagctggaac	caaaaccttt	gtcgggagat	ccgcagtctg	200580
cagtttgagc	acaggctctc	tggtttctgt	tctctgtcct	gtgtcgcata	ttgactagag	200640
gcagagaagc	atctgcaagg	ctgtgaccac	gctggctggg	gctctggccat	ctacattttc	200700
aacaggcaat	ctcaggagag	tatttccttt	taagaagcgc	agacttttgt	gcctgggcca	200760
cttctctact	tcccagaaca	ttgtgtgcca	agtggcaagt	tattaaccaa	gtgcttttga	200820
aaattaaact	ccttggtttg	cagagtagca	tgggagcatt	gagagggtgt	atgcctaaag	200880
gcctggttct	gctgctggca	gagctgacac	ttggctaaag	ggctggcatt	tctgagatga	200940
gcctcactag	atccgcgtct	cagagtctgc	aggagaaatc	agagagggga	gaaggtccag	201000
tggcctgttc	aggatgatct	tcctctgcac	ttaagggcgg	ctggtttgcc	cacgtagccc	201060
cagaacaaaa	cgagcctcgg	acgaagcccc	ctaaaggcag	taggagagac	tgagccttgg	201120
ctcttcagca	ggggtgggga	caagagcaag	aggcgggata	tcgcccggcc	ctttagagac	201180
acgtgcggtt	gtttccgtgt	ctgggagatc	acatgacccg	catcagctga	cccgtcacgg	201240
tgagctcag	cgctgggtgt	tcgcgctccc	cgccctgctg	cgccccggag	cgcaggaccc	201300
tgcggagggg	taagaaaacc	cccaggcttt	ctttcctttg	tcgctgggtc	gcgcagtcac	201360
ctgcacccta	cccccgctc	ctcgttcatc	ccagtcttcc	cggcctggca	ccccggaagc	201420
cactgcgagc	agggccgtgg	ccaggctcag	ccttgccgtg	cccccaggcg	gccaggacca	201480
aatgcccag	ggcggaagag	gggtcttaca	ggttcttaca	gcaggggtccg	agggctggct	201540
cccttcctca	ggacctgaca	tggaggagct	gctccggagc	gtggagagag	atctgaacat	201600
tgatgcccg	cagctggccc	tggcgccggg	gggcactcat	gtagtggccc	tagtgtccac	201660
gcgttggtctg	gctagtctcc	gggagcgccg	actgggaccc	tgtccccggg	ctgagggcct	201720
gggtgaagca	gaagtcagga	ctttactgca	acgttcggta	cagaggctgc	ccccaggctg	201780
gactcgagtg	gaggtgcatg	ggctgcggaa	acggagactg	tcctacccgc	tgggtggagg	201840
cgtgcccttt	gaggaggggt	cctgtagccc	tgaaactctc	actcggttca	tgcaggaggt	201900
ggctgcccag	aattaccgga	acctgtggcg	ccatgcatac	cacacttatg	gacagcctta	201960

cagccacagc	actgccccct	cagctctacc	tgccttagac	tctatacgac	aagctctcca	202020
gaggggtgat	ggatgcacct	tcttgccagt	gggtgaatcc	atcccatgtc	tatcaaagt	202080
cagggatggg	ccctgccccct	ctcggggcag	ccctgcctgc	cccagccttt	tgcgagctga	202140
ggctttgtctg	gagtcgccccg	agatgctcta	tgtggtacac	ccttatgtgc	aattctccct	202200
gcatgatgta	gttaccttca	gccctgcca	gctgaccaac	agccaagcca	aggtgctctt	202260
tcttctcttc	cgtgttctga	gggccatgga	tgcctgtcac	cgccaggggc	tggcctgtgg	202320
ggctctgtct	ttgcaccaca	ttgctgtaga	cgagaagcta	tgcagtgagc	tccggctgga	202380
cctgagcgct	tacgagatgc	cttccgagga	tgaaccaccag	gagggctctg	aagagaaaaa	202440
tgggacaggc	attaagtctg	aaaaagaggg	ggaagggaga	actgagtgtc	ccacttgcca	202500
gaaagaacct	cggggccttg	tgctagactg	ggcccatggc	cgaatcagca	acttccacta	202560
cctcatgcag	ctgaatcggt	tggcaggtcg	acggcagggg	gatcccaact	atcacccagt	202620
gctgccctgg	gtggtggact	ttaccacacc	ttatggggcg	ttccgagacc	ttcgtaaate	202680
caagttccga	ctcaacaagg	gagataagca	attggacttc	acctatgaga	tgacccggca	202740
ggcatttggt	gcaggtggtg	caggaagtgg	ggagccaccc	catgttcctc	accacatctc	202800
tgacgtgctc	tctgacatca	cgtactatgt	atacaaggcc	cgtcgcacac	cgcgctcggt	202860
gctctgtgga	catgtccgag	cgcagtggga	acccacagag	tatcctgcca	ccatggagcg	202920
gatgcagacc	cggtacacgg	atgagtgcac	accccatgtc	tacacggacc	cctctatctt	202980
ttgctctatc	caccctgaca	tggccgacct	ggatgtgccc	gcctgggtgca	gttctaacca	203040
ggaatttggt	gctgcccac	gagccctcct	ggagagctgg	gaggtgtccc	aagacctgca	203100
tactggatt	gatcttacct	ttggctacaa	actccagggc	aaagaagctg	tgaaggagaa	203160
gaatgtgtgt	ctgcacctgg	tggacgctca	cacccatctg	accagctatg	gcgtgggtaca	203220
gctatttgat	cagccacacc	cccaacgcct	ggctggatct	cctgccccgg	ccccgaacc	203280
tccactcatc	ccccggctgt	tggtccagcc	tattcgggag	gccacaggcc	aggagacat	203340
ttcaggacaa	ctataaaatg	gtgcggcgag	ctgtgtcgta	gaggccactc	catgtgagac	203400
tggctggact	agagataggc	ctgggacagg	agaagatgat	ttagaacagg	ctacagaagc	203460
tctggattcc	atctccctcc	ccgggaaagc	aggtgaccag	ccaggctctt	cctccagtca	203520
agcatcacct	ggcctgttgt	ctttttctgc	accctcgggg	tctcgaccag	gccgtaggag	203580
caaagctgcc	gggttggaac	ctggggaggg	tgaagagggc	aagattgtcc	ttccagaggg	203640
cttcagtcct	atacaggcct	tgggaagagct	ggagaagaag	ggtaacttcc	tggccaaagg	203700
cctaggggag	cagttggagg	agcctgaaaa	gcctcacgcc	cagccacctg	tgcacctgca	203760
gagcctcttc	catcgagaca	tgcaggtcct	gggtgtcctg	ttggctgaga	tgggtgtttgc	203820
caccagggtc	cggatactgc	agcctgatgc	acctttgtgg	gtacgctttg	aggctgttcg	203880
gggtctctgc	atacgccact	ccaaggacat	ccccgtgtct	ctgcagcctg	tgctagacac	203940
actcctacag	ctgagcggac	ccaaaagtcc	catggtgtcg	aagaagggca	agctagaccc	204000
actgtttgag	tataggccgg	tttcccaggg	attaccccca	cccagcccag	cccagctcct	204060
cagccccctc	agctccgtgg	tccccttccc	tccatacttc	ccagcactgc	acaagttcat	204120
tcttttatat	caggccccgg	gtgtggaggga	tgaggtccag	ggtcggggag	tggcgtttgc	204180
tctgtggcag	cagctgggtg	cggtgttaaa	tgacatcact	cccaggggct	tagagatcct	204240
cctgccccct	gtgctgtcgc	tcatgtctga	ggagcacacg	gctgtgtaca	cagcctggta	204300
cctatttgaa	cccgttgcca	aggccctggg	ccccaaaaat	gccaacaagt	acctcctgaa	204360
gcctctcatc	ggtgcctatg	agagccccct	ccgcctgcat	ggccgcttct	acctgtacac	204420
cgaactgttt	gtggcccagt	tggtgggtgc	gctgggcttg	caggccttcc	tcacccacct	204480
gctgccccct	gtctccaggg	tactggctgg	gggtggaggt	tcccaggagg	agggcaaagg	204540
cctgggtcggg	accactgagg	atgaggaaag	tgagctcccg	gtgtccgggc	ctggctcctg	204600
tgcctttggg	gaagagattc	agatggatgg	gcagccggct	gcttccctcag	gactggggct	204660
cccagactac	aggtcggggc	tcagcttcca	tgaccaggcc	gacctgccgg	acacggagga	204720
cttccaagct	ggactctacg	tggctgaatc	tccacagccc	caggaggctg	aggccgtgag	204780
cctggggccag	ctgagtgata	agagcagtag	cagcgaagcc	tcccagggcg	aggagagggg	204840
tggggatgat	ggcgggtgcc	ctgcggacaa	gaacagcgct	aagtcagggg	acagcagcca	204900
ggacttgaa	cagagcgaag	gctctgagga	agaggaggag	gaggaaggct	gtgtgggtgt	204960
ggaggaggac	caggaggatg	aagtcacggg	aacatccgag	ctcactctgt	ctgacacgat	205020
gctgtccatg	gagacgggtg	tggctcctgg	tgatggggaga	gacagagaag	aggagagagga	205080
gccgctgaca	gagcagacag	aaggcaaaaga	acaaaagatc	ctccttggtg	agcccgtggg	205140
ctgagggggc	atgggtcagg	tgcttttctc	tcaggctctc	atatgctggg	tgtgggtcca	205200
accagatcca	ctgtagcacg	cacagccaca	gtcagacaca	gtgcatggaa	tgtggaagtg	205260
ctgtgtgtga	gtggaaagtg	gggcttagat	ttagctttca	ggagacagaa	agctccttta	205320
aaagccatac	cttgggctga	ggctggggag	ggagttgagt	ggtagagcac	ttgtctggta	205380
tactcgaggg	ccctgggtgt	cttatctcta	gccccagaag	aagtattaa	aaataaaagc	205440
aagtgggtgt	tgagatgtga	atggagccag	aactggccgg	aacagtcggg	tggagagtgg	205500
aagagtgttc	cagacagggga	acagtgtgtg	tgtacctctg	aggctctcat	ggttccatca	205560
gagaggcagg	gaaaggctaa	aatggttttc	ttaagagagt	ccagaagggc	tgggcttggt	205620
ggtgcaggcc	tttaatccca	gcactcagga	tgacagggca	ggcggttttt	tgagttcgag	205680
gccagcctgg	tctacaaagt	gagttccagg	acagccaggg	ctatacagag	aaaccctgtc	205740
tcgaaaaaaa	aaaaaaaaaa	aaagagtgtg	gaaggggtgga	agccagggac	aagctctgtac	205800

aagaaggaac	ttgggagcat	tgccgaaagg	atgacctctc	tgcaggctct	gccccaggag	205860
ccagtttctg	gggaccttga	ccatggctag	gtgaatggac	ccaggatggg	atggtcaggc	205920
ttgctagcag	agccacagcc	gagttggctg	ggtgggggtg	ggtgggggtg	gaagggtgag	205980
ttatctgatg	agctcaggac	cttttccctgc	cctgcagata	cagctgcaaa	gatggtaccg	206040
tggctgtctg	ccaagcttgg	ccccacagta	gcctctcgcc	atgtggcccc	gaacctgctg	206100
cgcctgctga	catcttggtt	tgttggttaag	gtctgtgggt	agtgtctggag	accagggttc	206160
ccagccaggc	ttctgcccac	ccttagccct	ctctaggcga	ctccttccct	aacttcccag	206220
cactccctga	gcagggcctg	ggtctcaccc	attaagctgg	gttttcttgg	gtaagtgggg	206280
aagagccccag	tattgaatga	atagaagcca	ccccacagtc	tcagaaggcc	ggcttccctc	206340
ctgccctcca	ctggcttctc	aacgtctctg	cccttccctg	gtagggcccc	ctcgacagca	206400
gttcaccgtg	agcagtgatg	acaccctccc	actgaatgcc	ggcaacatct	accgaagag	206460
gccagtccta	ggtgacatcg	tgtcggggcc	tgtgtctcagc	tgccctccctc	acattgccta	206520
cctgtatgga	gaaccctgtt	tcacctacca	gtacctgccc	tacatcagct	acctggtcag	206580
tccctgggtt	gtcaaaccct	ggcttggggg	tgggggcaag	gatccaagga	ccagccccag	206640
gtcttggggg	ttccaggagg	tctgtggggg	gacctgtccc	tccctcatct	attctgtggt	206700
tctaggtagc	cccaggggagc	aactcaaacc	ccagccgact	gaacagccgc	aaggaggccc	206760
ggctgtgggc	agcggtgaca	ctgacgcaga	aaatcccttg	atacctctct	gacacgacc	206820
tcatggacat	tctgccccgc	attagccacg	aggtcttgct	gcctgtgctt	ggcttccctc	206880
cctccttctg	cacagggtag	gcccctgctg	cttgggagag	ccacctggct	gagggggccc	206940
ccaggaaggg	ctaggaagct	cagggagaag	cagataccgg	cctgagtcac	ggttctgatg	207000
ttgggggtag	tggcacaggt	ctttcattcc	agcaccaga	ggagggcaag	tttctgtgag	207060
tctgagacta	gcctgggtcta	cagagagagc	tccaggctat	ctaaggctcc	atagtaagac	207120
tctgacttaa	gaaaagagtc	gtggttcatt	ctgggttggt	ggtgtggctt	ggtgatggga	207180
cactttccca	gcctgcagga	ggagctatgc	ttgagttcca	gcccttcaga	aaaacaaaaa	207240
tgggggctgg	aaagaatagc	tcagggttta	agagcactgg	ttgctcttcc	agaggatcca	207300
ggttagattc	ccagctgcca	catggtagct	cataaccatc	cggcagttct	atggaacctg	207360
ccaccctcct	tcgggtctctg	tgggcaactg	aaacatgtgc	acagacatac	atgcaggcag	207420
aaaaaacacc	catacacata	aaattagacc	aaaaaagttc	atgttctctc	ctacctgtag	207480
ctctgactaa	gtacactgct	ttccctgtgc	ctcagtttcc	tccctgggtc	tggactgatc	207540
agccttacat	gcagctcctg	ttatttgaag	ttcctggtaa	attgggtcaag	tccttcaggg	207600
aagggtctgg	aactcttgca	ctttgattct	aggttcccca	gtggggccca	gccccggact	207660
gtcctatgct	tgaaaacat	cagtctcatc	gccctcatct	gcttgcgcat	cgggcaggag	207720
atggtccagc	agcacctgag	tgagccagtg	gccaccttct	tccaagtctt	ctctcatctg	207780
catgagcttc	ggcagcaggt	aggcaggcag	cttctgggct	gggtgggcca	ggccaggcca	207840
ggccaggcca	gggcagtgga	cccactgaat	ctgtggtctt	cctacccgca	ggatctgcca	207900
ctggatccta	agggtgttac	tgaggggccag	ctgccagagg	cgaccttctc	tgatgggcag	207960
cgacgaccag	tggacccac	cctgtgggaa	gagctgcaga	aggtgttcac	cctggaaatg	208020
gcgtacacaa	tctacgtacc	tttctcctgc	ctgttgggta	ttgcccatac	cgttcccttg	208080
cacagagttg	gtgactacat	ctcttccctg	gggtgggccc	cgatgctttc	acctccagag	208140
tcagcaatgg	aatcttttta	tttttatttt	gacatggggg	ctcatttagc	ccaggctgac	208200
ctttaactcc	agctccttcc	agcttccacc	gtctcctgtt	ggcattgtag	tcatgtggca	208260
ttgctcaggc	ttcttncatg	ttcttatttt	taaatgacct	gtgtgtgtgt	gtgatctctg	208320
tgtgagtgtg	gaggtgacag	aataacagtt	ggggggctag	cagatgcctt	gcctgatgag	208380
catctctcta	gatccagttt	ttgggtttgt	ggggttttat	gtgtgtgttt	gtttgtctgt	208440
ttttgtagac	agggtctctc	tgtgtagcct	ggccatcctg	gaactcattc	agtagaccac	208500
gctggccttg	agctcacaga	gattcacctg	cctctgcctc	ccagtgtctg	gattaaaggc	208560
gtgtaccact	cctgcctggc	tttgtttttg	ttaaccacca	tctcctgccc	tcagcatctg	208620
cctccccctg	gctgggatta	cagggtgtgtg	ctatcacacc	cagctaacag	tggatttaaa	208680
cgtaggaatt	ttaggatcag	agtgaaccaga	tttggtccta	gggccaatt	tccacagtga	208740
ttatctatct	tagttaggat	ctctgttgaa	aatcatggtg	gaacatcatt	accaaagcga	208800
acttggggag	gaaaagggtt	attttgtctg	acaactctca	ggtcaccaag	ggaagtcagg	208860
gcagggaactc	gaggcagaag	ctgaagcaaa	agccatggaa	gaactctggc	ttgttctctc	208920
tggcttgctc	agtctgggtg	acccctccc	caccacctc	cccacaatgg	tttctctgct	208980
tatgcctggc	tgtcctagaa	ctcactctgt	agactaggct	ggcctcaaac	tcaagagatc	209040
cccctgctc	tgcactctca	gtgctaagat	taaaggcggg	tgccatcacc	cctgccccag	209100
gggtggcact	accactgta	aattgggtccc	tcccatact	agttgtttaa	taagaaaact	209160
cctccatagg	ccaatctggt	gggggaattt	tctcagttga	gggtttctct	tctcaaatga	209220
ctgtagctga	tgccaaattg	ataaaacaaa	tctcaaacca	ccaccaccaa	caacaataaa	209280
accaaacaaa	ccaaacaact	aaccaagaca	gtgacttata	aagagaatct	gaacattttc	209340
cagcaggaaa	ggctcaggag	ctggccattc	aagtctgggg	aacagaatgt	aggggaatat	209400
gatggtctcc	agaagctacc	tgcaaaggaa	tgaacagctt	gctgggtttt	gtgggttccc	209460
ttatgggtag	ggcgtgtac	tgggcttctc	cttgagttag	atgggcccac	ctgtgattgg	209520
gaatattttg	ctcctacaga	attgtaagtt	cccagaggca	ggacacatct	gtcttattct	209580
tcattgtgtg	tctgatgcta	gaatgggtgc	tggcatacac	gtgtgtgtct	ctatagagac	209640

agcactcatg	tctacgtatc	gataaaggaa	gctgttttgg	ggggaggaaa	caggcttaca	209700
gacgagaact	taataaccca	gagtagccca	gtcagtagct	tgccttggtt	tctgtttgtt	209760
ctaagctctg	ggtagatagc	taccttgcca	tcttccctga	tcttagaact	ttccccactc	209820
ccctgtaggt	gacatcatcc	ggaaaatcat	ccccaaccat	gagttgggtc	gggagctggc	209880
agggtcttat	ctggaaagca	tgagcccagag	ctctcgaaac	ccagccagca	tggaaaccac	209940
catggctagt	gccggccctg	aatgggaccc	tcagagtggg	agctgtctcc	aggacgatgg	210000
ccactcaggg	acctttggga	gtgtcctggt	tggaaatcgc	atccagatcc	ctgactctca	210060
gccccagagt	cctgggccac	tgggctccct	ctctggagtg	ggtagtagcg	gaggcctcag	210120
caacaggaat	gaagacaacg	ccctgaagcg	ggagctgcct	cggagtgcgc	atgggctgag	210180
cgggaactgg	ctggcgctact	ggcagtagca	gatcgggtgt	agccagcagg	atgcccactt	210240
ccacttccac	cagatccgcc	tgcagagctt	cccaggcac	acggggggcc	tcaaactcgt	210300
ggccgcccctg	agcagtgaag	acttctttct	gagtgggcag	aaggaccgga	ctgtgcgcct	210360
ctggccgctg	tacaactatg	gggacgggac	caatgagacg	gcttcccgc	tcatctatgc	210420
ccagcaccgc	aaaagcgtct	tctacgtggg	ccagcttgag	gccccgcagt	atgtggtgag	210480
ctgtgatggg	gcagtgcacg	tctgggaccc	cttcacaggt	gagcggggccc	aggtgaggcc	210540
tgttcgacgg	ctgctttact	gtgccttagc	caggcctctg	ggaacgggac	ctagtgcgaa	210600
acgtacaatg	gcgtattttg	acggggaaga	ttcagtggag	caggaagaga	agaagagtca	210660
ggacttagaa	tctgtgggac	ccaagtttga	atccactccc	ccaacttacc	agcaatcggc	210720
tcagtgtctg	caggcgtctg	ccttctacct	gtaagaacca	aaaattttaga	agattccacg	210780
agtatggctt	tggcttcttg	tacgacgtca	cctgtcgtcg	ttgtaaagag	aagtatcgag	210840
tggaggaggg	tcagggcaga	cggaggtcgc	agctagttag	agcatgctat	gtgaagagag	210900
cagactgttc	tggggctgga	cccttgactt	cactgtggaa	gcagcaagat	gagaaagccc	210960
ttgagatttg	ttttctgagg	gtcactgggg	aatgggatgc	aggtgtgggg	tgagttggag	211020
ttgaagtag	ccagggtctt	ttgatagcca	ctaagtcctc	agatgtgtcc	tttttcagga	211080
aagacccttc	gcacagtggg	tccttcagac	agccgggtgc	ccctgacggc	tgtggctgtc	211140
atgcctgccc	cacacaccag	catcaccatg	gccagctccg	actccactct	gcgctttgtg	211200
gactgcagga	agccaggctt	gcaggtcagg	aggggtgcag	ttcctgggct	actgggggtc	211260
tctaggtacc	agtcaggaaa	gacactcagg	ggactccacc	aggaacgctg	cagtgcacag	211320
cagccctgtg	tgggtggggc	gctggcacgg	atggggcttt	tctcttccgg	ggatggagtg	211380
ggagggtcag	gcctactggt	ttcgtgggcc	tgaatggggt	gagctgcagt	aggggtgggtg	211440
gcagtgatgg	atggcgacgg	gcacttgaac	acaatctcct	cctatagcat	gagttccgac	211500
tgggtggagg	gtggaacctt	gggcttggtc	gctcgttggc	cgtcagcccc	agtggccgga	211560
gtgttgtggc	tggcttctcc	tcgggcttca	tgggtgtcct	agatacccg	acgggcctgg	211620
ttctacgagg	ctggccagcc	catgaagggg	acattctaca	gatcaagggt	actgactgcc	211680
tgaggtccta	tcctttcatt	tctacttagg	gcctgggtct	ggagaggaca	ggtttatgct	211740
ggtgtccctt	ataactactc	ggggacattc	agtgggggtg	gaaaatggcc	ctcgtaggcc	211800
agctcaggaa	ccagctgcac	aggaggcagg	ctaggggcag	gaatcagggc	tagaactgac	211860
cctgatgttc	cacagcgatg	ttctaattag	taacccttct	ccataattgt	cttgcttggg	211920
ggatcagggg	tcacgcccctg	tccgtgaccc	agttcagggt	aaataaagcc	aggaggctgt	211980
ttactgcctg	gagaccactg	agcagagtcc	atgccccctg	ctgggctgtc	ctgatggggg	212040
gcaggaacag	gcgcaggcct	gcgcacgtg	ttcctgcctc	ctatattcaa	tcatagacct	212100
cagagctcag	caggttcttg	gaggggagaa	atagggctgc	ttgtgggagg	atttctccct	212160
gcagtgggaa	ctctcctccc	ccgcccgtca	ctaggaggtg	aaagacaggc	actgttgcct	212220
acagggaagg	caggctggcc	ccagctctat	cgagaccccc	aggggaccct	gggtctcagt	212280
gtctctaaat	cccaacattc	taagaaagtg	tcaggatggc	tctgggggtc	tcctgggtgt	212340
tagtcccagc	tctcggagtc	ttctctgagc	accagttctt	ttcctatggg	aagtaaggac	212400
atgccagggtg	ttctttgaga	ggacctgagt	ttgggtctca	gcactgtcta	gctctggctc	212460
caggggggttc	aacacccttt	atggcttctg	tggacacata	ttcttatgtg	gcctgcacgc	212520
acacacacga	acacaaataa	aaataaatgt	taaaagaaga	cagcggcacc	ttgtacctca	212580
catgttagta	cgattggatg	tggcagtgcc	tcacagaatc	tgtgggactt	tattttattta	212640
tttatttttt	ggtcttataaa	ttttaaaaga	ttggttttagg	agtgggtgga	cacaccatta	212700
atcgcaacac	tcaggagcag	aggcagggtg	atctctatgg	gtttgaggcc	agcctggtct	212760
acagagcaag	tttcagggca	gccaagggtta	cacagagaaa	ctctttctca	aaaaataaaa	212820
acaaaatatt	taaaagattt	acttacttct	tatttgagct	aggatctccc	tatttagccct	212880
ggctgtcctg	gaactcactg	tatagaccag	gctggcacct	taaactcacg	aagatcctcc	212940
tgctttctgc	tcctaagtgc	tgagattaaa	gtagtgttat	accatgcccc	actattttct	213000
ttatagattt	ggtgttttgc	ctgcttgtgt	atatatgcac	taccttcatg	cagtgttaat	213060
aggggtcaga	ggcgagtatc	agctcttcc	ggaactagag	ttatggaagg	ttgggagtca	213120
ccatgctggg	actgggtcat	ttgcaagagt	tacaagtact	tctgagccat	ctccagcccc	213180
ctagagtttt	tttccccctt	ggctgtcctg	aagtagaatc	tgcttctgtt	ttgttttgtt	213240
tttcgagaca	gggtttctct	acataggcct	ggctgtcctg	gaactcactc	tgtagaccag	213300
gctggcctcg	aactcagaaa	tccgcctgcc	tctgtgcttc	agaatactgg	aattaaaggt	213360
gtgcgcaccc	acgcctggct	cagaatctgt	ttgttaatga	gagtaatagt	tacaggtttt	213420
ttgttttgtt	tttttctttt	ctttttttgt	ttttgttttt	tggctcattt	gtttttttgt	213480


```

ttttgagaca ggtctcactc tgaaccccta ggtggcctgg agcttgctat gtagaacaca 213540
ctgacttttaa acttgttttc tgagtgtctg atttatgggc ttgtgctatt ttgccagacc 213600
tctgatgggt gttaataaca atattattta gcttttcttt tggagatagg ctctcactgt 213660
atatcaccca gacagtggct ggtctggaaa tcaactgtga gcccaggctg acctgaatt 213720
cacagagatc tgcctcccga gttcagaaat taaaagcact ctgggatggg tttggagttt 213780
ggtgagtacc caagcctcca ttgatgctat ctgtccctcc cgctctctgc aggctgtaga 213840
gggcagcgtg ctcatcagct cctcttccga ccattccttg actgtttgga aggagctgga 213900
acagaagccc acgcaccact acaagtcagc gtccgaccca atccacacct ttgacctgta 213960
cggcagcgag gtggtcaccg gcactgtagc caacaagatt ggtgtctgtt ccctgcttga 214020
gccaccctct caggccacca caaagctcag ttccgagaac ttccgtggca cgctcactag 214080
tctggctttg ctgcccacga aacgccacct cctgtctggc tcggacaatg gcatcatccg 214140
cctcctggca tagggccagc caggagtgtg ctgagggcag ggcgagatga catctctcag 214200
ggcccgtctc tcattcttga tctcgaagcc gattcttcta ggcaagcccc aggctctggc 214260
taccacatg gcctgctgtc tgggattgca cagctcctga atctccaaag ccttgaagtg 214320
gcttcatgaa actcgggaga tactgttcct aaccagcaag aattggggca aggaaagcac 214380
tgtgatcccc attgctcccc agttctgcct tctggattca catggggaca gggcagctcc 214440
aggaaatgaa aggagtggg cctttgctca gccagcttcc tctagccacg ctctccttag 214500
ctctgtttct cccttgggta ggaaactgct cctgtctagg gttctgatgg tactgggact 214560
ccaggtctag cagggtggc caggacctac gacttccagg gcttggctctg ggggttttagc 214620
attcattcag ccaggtcttc agtatgggac cagaaaaaag gggatgtgag aacagggcta 214680
gggaaggggt tatatgggcc cagctgggtc aggaatgaat ccatgccttg ccttgggtacc 214740
cctaaccaca gcgtttgtgc cttcagccgg ggaggcagcc cttgggacca gcatccctag 214800
ggacaggagg cagcgggaat catctctgta tctcgggttc tgcccagggg atgggcagac 214860
tctgccatct cttgagtgtt cgtttgaga agcctgagat gtggcccctg ctgccttctc 214920
actagtgtca gtctatgtaa ataaggtcaa taaattcttt ggaagagcca cggagctgag 214980
tgaggctgtg ttgtgttttg ctttgcttag gctgggtcctc ggcagctctg cctcagcctc 215040
ccaaggagct ggggaactgg tatatgtcac tgtatatgtc actgtgcctg gcttatggct 215100
tggcttggct ttttttcaga tggctcgaag tgcctcaggt tggccttgat cttgggatga 215160
ccttcctgct tgaacacagag tagtgggctt ataggcatga cccaccagggt ccaattttta 215220
ttttttaaag gcattgattt ttatacgtgt atggttgttt tgcccacttg tacatatgca 215280
caccatactt gtgtctggtc cctgcggagg tcagaagagg gcatcgggat cacctggaac 215340
cgaagttaat gaatggttat gagccacatc tcatgctga agattgaacc tggatccttt 215400
gcaagagcag ccagtgttct taccactga gccatctcta agccccacac ccagcttctt 215460
ttgatacaag gtctggtagc tcaaaactga tatgcagccg aggaggttga cctgggtattc 215520
cctacctacc ctcttctctc taccttccaa gtgctgatat tatacatagg catggatagt 215580
catgcccacc agtttgcctt gatggcacca gagtcaggaa agtccaaacc tggtagttgc 215640
aaacacagca agagggtaga ggcagccatt gtcctctggc tgccttggat acagagcttc 215700
tgggttgggt ggccttgggt cagttttccg aatgggtcac ccttggggaa agggaacact 215760
gctgaagagg tgggaccctg ggagggccgg cctccagctg ggtctctcca gccctgcct 215820
tggaacctag gctggaggga gccaacagg atcctggact tgctacagtt aggtgaacag 215880
gctcctgcag cctccccttc ccttgggtag ctgtgggtgg ggtgggtggtg gtgggtggtg 215940
tggtggtggt ggtggtggtg gtgggggggg gngnngnt 215980

```

<210> 17
 <211> 473
 <212> PRT
 <213> Mus sp.

<400> 17
 Met Lys Arg Ala Ser Ser Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
 1 5 10 15
 Trp Leu Gln Ala Trp Arg Val Ala Thr Pro Cys Pro Gly Ala Cys Val
 20 25 30
 Cys Tyr Asn Glu Pro Lys Val Thr Thr Ser Cys Pro Gln Gln Gly Leu
 35 40 45
 Gln Ala Val Pro Thr Gly Ile Pro Ala Ser Ser Gln Arg Ile Phe Leu
 50 55 60
 His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Gln Ser Cys
 65 70 75 80

Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Ala Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Thr Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu His Val Val Asp Pro Thr Thr Phe His Gly
 115 120 125
 Leu Gly His Leu His Thr Leu His Leu Asp Arg Cys Gly Leu Arg Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Asn Leu Gln Ala Leu Pro Asp Asn Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Pro Ser
 180 185 190
 Val Pro Glu His Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu
 195 200 205
 Leu His Gln Asn His Val Ala Arg Val His Pro His Ala Phe Arg Asp
 210 215 220
 Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Met
 225 230 235 240
 Leu Pro Ala Glu Val Leu Met Pro Leu Arg Ser Leu Gln Tyr Leu Arg
 245 250 255
 Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp
 260 265 270
 Ala Trp Leu Gln Lys Phe Arg Gly Ser Ser Ser Glu Val Pro Cys Asn
 275 280 285
 Leu Pro Gln Arg Leu Ala Asp Arg Asp Leu Lys Arg Leu Ala Ala Ser
 290 295 300
 Asp Leu Glu Gly Cys Ala Val Ala Ser Gly Pro Phe Arg Pro Ile Gln
 305 310 315 320
 Thr Ser Gln Leu Thr Asp Glu Glu Leu Leu Ser Leu Pro Lys Cys Cys
 325 330 335
 Gln Pro Asp Ala Ala Asp Lys Ala Ser Val Leu Glu Pro Gly Arg Pro
 340 345 350
 Ala Ser Ala Gly Asn Ala Leu Lys Gly Arg Val Pro Pro Gly Asp Thr
 355 360 365
 Pro Pro Gly Asn Gly Ser Gly Pro Arg His Ile Asn Asp Ser Pro Phe
 370 375 380
 Gly Thr Leu Pro Ser Ser Ala Glu Pro Pro Leu Thr Ala Leu Arg Pro
 385 390 395 400
 Gly Gly Ser Glu Pro Pro Gly Leu Pro Thr Thr Gly Pro Arg Arg Arg
 405 410 415
 Pro Gly Cys Ser Arg Lys Asn Arg Thr Arg Ser His Cys Arg Leu Gly

	420						425						430				
Gln	Ala	Gly	Ser	Gly	Ala	Ser	Gly	Thr	Gly	Asp	Ala	Glu	Gly	Ser	Gly		
		435					440					445					
Ala	Leu	Pro	Ala	Leu	Ala	Cys	Ser	Leu	Ala	Pro	Leu	Gly	Leu	Ala	Leu		
	450					455					460						
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys									
465					470												

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/029059 A3(51) International Patent Classification⁷: C12N 15/12,
C07K 14/705, 16/28, C12N 15/62, G01N 33/68, A61K
38/1706437 (US). CATE, Richard, L. [US/US]; 40B Nichols
Road, Cohasset, MA 02025 (US). SAH, Dinah, W., Y.
[US/US]; 4 Longfellow Place, Apt. 2608, Boston, MA
02114 (US).

(21) International Application Number: PCT/US01/31488

(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251
Avenue of the Americas, New York, NY 10020 (US).

(22) International Filing Date: 6 October 2001 (06.10.2001)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
60/238,361 6 October 2000 (06.10.2000) US(71) Applicants (*for all designated States except US*): YALE
UNIVERSITY [US/US]; 541 College Street, New Haven,
CT 06520 (US). BIOGEN, INC. [US/US]; 14 Cambridge
Center, Cambridge, MA 02142 (US).(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): STRITTMATTER,
Stephen, M. [US/US]; 96 Tulip Tree Drive, Guilford, CT

[Continued on next page]

(54) Title: NOGO RECEPTOR HOMOLOGS

	1				50	
Human NOGO-R1	-----	-----	MKRASA	GGSRLLAWVL	WLQAWQVAAP	CPGACVCYNE
Murine NOGO-R1	-----	-----	MKRASS	GGSRLLAWVL	WLQAWRVATP	CPGACVCYNE
Murine NOGO-R3	MSWQSGTTVT	QSPVQAAQVS	GCCVELLLLL	LAGEPLPLGG	CPRDCVCYPA	
Human NOGO-R3	-----	-----	-----	-----	-----	-----
Human NOGO-R2	-----	-----	MLPGLRRLQ	APASACLLM	LLALPLAAPS	CPMLCTCYSS
Consensus	-----	-----	-----	-----	-----	CP--C-CY--
	51	LRR NT		LRR 1		100
Human NOGO-R1	PKVTTSCPQQ	GLQAVPVGIP	AASQRIFLHG	NRISHVPAAS	FRACNLITIL	
Murine NOGO-R1	PKVTTSCPQQ	GLQAVPTGIP	ASSQRIFLHG	NRISHVPAAS	FQSCNLITIL	
Murine NOGO-R3	P.MTVSCQAH	NFAAIEGIP	EDSERIFLQN	NRITFLQGGH	FSP. AMVTL	
Human NOGO-R3	-----	-----	EGIP	VDSEVFLQN	NRIGLLQPGH	FSP. AMVTL
Human NOGO-R2	P. PTVSCQAN	NFSSVPLSLP	PSTQRLFLQN	NLIRTLRPGT	FGS. NLLTL	
Consensus	P--T-SC--	-----P--P	-----R-FL--	N-I-----	F-----	---L
	101	LRR 2		LRR 3		150
Human NOGO-R1	WLHSNVLARI	DAAFTGLAL	LEQLDLSDNA	QLRSVDPATF	HGLGRHLTLH	
Murine NOGO-R1	WLHSNALARI	DAAFTGLTL	LEQLDLSDNA	QLHVVDPTTF	HGLGHLTLH	
Murine NOGO-R3	WIYSNNITFI	APNTFEGFVH	LEELDLDGDN	QLRTLAPETF	OGLVKLHALY	
Human NOGO-R3	WIYSNNITYI	HPSTFEGFVH	LEELDLDGDN	QLRTLAPETF	OGLVKLHALY	
Human NOGO-R2	WLFSNNLSTI	YPGTFRHLQA	LEELDLDGDN	HLRSLEPDTF	OGLERLQSLH	
Consensus	W--SN----I	----F----	LE-LDL-DN-	-L-----P-TF	-GL--L--L-	
	151	LRR 4		LRR 5		200
Human NOGO-R1	LDRCGLQELG	PGLFRGLAAL	QYLYLQDNAL	QALPDDTFRD	LGNLTHLFLH	
Murine NOGO-R1	LDRCGLRELG	PGLFRGLAAL	QYLYLQDNAL	QALPDNTFRD	LGNLTHLFLH	
Murine NOGO-R3	LYKCGLSALP	AGIFGGLHSL	QYLYLQDNHI	EYLDQDIFVD	LVNLSHLFLH	
Human NOGO-R3	LYKCGLSALP	AGVFGGLHSL	QYLYLQDNHI	EYLDQDIFVD	LVNLSHLFLH	
Human NOGO-R2	LYRCQLSSLP	GNIFRGLVSL	QYLYLQENSL	LHLQDDLFD	LANLSHLFLH	
Consensus	L--C-L--L-	---F-GL--L	QYLYLQ-N--	--L-D--F-D	L-NL-HLFLH	

(57) Abstract: The invention relates generally to genes that encode proteins that inhibit axonal growth. The invention relates specifically to genes encoding NgR protein homologs in humans and mice. The invention also includes compositions and methods for modulating the expression and activity of Nogo and the NgR proteins. Specifically, the invention includes peptides, proteins and antibodies that block Nogo-mediated inhibition of axonal extension. The compositions and methods of the invention are useful in the treatment of cranial or cerebral trauma, spinal cord injury, stroke or a demyelinating disease.

WO 02/029059 A3



Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(88) Date of publication of the international search report:

23 January 2003

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/31488

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/705 C07K16/28 C12N15/62 G01N33/68
A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBL, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL NUCLEIC ACIDS 'Online! 21 January 2000 (2000-01-21) "M. musculus chromosome 11 clone RP23-113D4" retrieved from EMBL Database accession no. AC021768 XP002217414 compare translation product of complementary strand of nt 153209-154529 with aa 21-461 of seq. 4 and aa 1-392 of seq. 14 --- -/--	1,2,5-30



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the International filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

18 October 2002

Date of mailing of the international search report

08/11/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Galli, I

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/31488

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL NUCLEIC ACIDS 'Online! 15 November 1999 (1999-11-15) "Homo sapiens clone RP11-10E2" retrieved from EMBL Database accession no. AC013606 XP002217419 compare aa 168-420 of seq. 2 with the translation product of nt 138578-139331.</p>	1-4,6-30
X	<p>DATABASE EMBL PROTEINS 'Online! 1 May 2000 (2000-05-01) CORBY N.: "New leucine-rich protein" retrieved from EMBL Database accession no. Q9UGS3 XP002217415 compare aa 352-399 with seq. 11,19</p>	1,11,15
A	<p>DATABASE EMBL PROTEINS 'Online! 1 October 2000 (2000-10-01) OSADA N. ET AL.: "Unnamed protein product" retrieved from EMBL Database accession no. Q9N0E3 XP002217416 compare with sequences 2,4,5,11,17,18,19</p>	1-30
A	<p>WO 00 31235 A (CHEN MAIO S ;SCHWAB MARTIN E (CH)) 2 June 2000 (2000-06-02) the whole document</p>	1-30
P,A	<p>WO 01 51520 A (UNIV YALE ;STRITTMATTER STEPHEN M (US)) 19 July 2001 (2001-07-19) NgR1: compare seq. 1,3 with seq. 5,17 of present application, respectively</p>	1-30

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims 1,2,6,7,9-13,15,16,18-30 relate to an extremely large number of possible compounds. It is not possible to determine all the possibilities, as any given X can be any amino acid or even a gap. In fact, the claims contain so many options, variables, possible permutations and provisos that a lack of clarity (and/or conciseness) within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible. Consequently, the search has been carried out for those parts of the application which do appear to be clear (and/or concise), namely seq. ID 2, 4 and 14.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/31488

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 27, 29 and, as far as methods in vivo are concerned, claims 26 and 28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/31488

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0031235	A	02-06-2000	AU 1469200 A	13-06-2000
			CN 1354755 T	19-06-2002
			CZ 20011608 A3	17-10-2001
			EP 1124846 A2	22-08-2001
			NO 20012223 A	02-07-2001
			SK 6222001 A3	03-12-2001
			WO 0031235 A2	02-06-2000
WO 0151520	A	19-07-2001	AU 2940101 A	24-07-2001
			EP 1248803 A2	16-10-2002
			WO 0151520 A2	19-07-2001
			US 2002012965 A1	31-01-2002
			US 2002077295 A1	20-06-2002

Form PCT/ISA/210 (patent family annex) (July 1992)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
11 April 2002 (11.04.2002)

PCT

(10) International Publication Number
WO 02/029059 A3(51) International Patent Classification⁷: C12N 15/12,
C07K 14/705, 16/28, C12N 15/62, G01N 33/68, A61K
38/17

(21) International Application Number: PCT/US01/31488

(22) International Filing Date: 6 October 2001 (06.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/238,361 6 October 2000 (06.10.2000) US(71) Applicants (for all designated States except US): YALE
UNIVERSITY [US/US]; 2 Whitney Avenue, New Haven,
CT 06511 (US). BIOGEN, INC. [US/US]; 14 Cambridge
Center, Cambridge, MA 02142 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): STRITTMATTER,

Stephen, M. [US/US]; 96 Tulip Tree Drive, Guilford, CT
06437 (US). CATE, Richard, L. [US/US]; 40B Nichols
Road, Cohasset, MA 02025 (US). SAH, Dinah, W., Y.
[US/US]; 4 Longfellow Place, Apt. 2608, Boston, MA
02114 (US).(74) Agents: HALEY, James, F. et al.; Fish & Neave, 1251
Avenue of the Americas, New York, NY 10020 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,

[Continued on next page]

(54) Title: NOGO RECEPTOR HOMOLOGS

	1		50
Human NOGO-R1	-----MKRASA	GGSRLLAWVL WLQAWQVAAP	CPGACVCYNE
Murine NOGO-R1	-----MKRASS	GGSRLLAWVL WLQAWRVATP	CPGACVCYNE
Murine NOGO-R3	MSWQSGTTVT	QSPVQAAQVS GCCVELLLL	LAGEPLGGG CPRDCVCYPA
Human NOGO-R3	-----	-----	-----
Human NOGO-R2	MLPGLRELLQ	APASACLLM LLALPLAAPS	CMILCTCYSS
Consensus	-----	-----	CP--C-CY--
	51	LRR NT	LRR 1
Human NOGO-R1	PKVTTSCPQO GLQAVPVGIP	AASORIFLHG NRISHVPAAS	FRACRLTIL
Murine NOGO-R1	PKVTTSCPQO GLQAVPTGIP	ASSQRIFLHG NRISHVPAAS	FOSCRILTIL
Murine NOGO-R3	P.MTVSCQAH NFPAIPEGIP	EDSERIFLQN NRITFLQGGH	FSP.AMVTL
Human NOGO-R3	-----EGIP	VDSEVFLOQ NRIGLIQPGH	FSP.AMVTL
Human NOGO-R2	P.PTVSCQAN NFSSVPLSLP	PSTORLFLQN NLIRTLRPGT	FGS.NLLTL
Consensus	P--T-SC--	-----P--P	-----R-FL--N-I-----F-----L--L--
	101	LRR 2	LRR 3
Human NOGO-R1	WLHSNVLARI DAAAFGLA	LEQLDLSDNA QLRSDVPATF	HGLGRHLTLH
Murine NOGO-R1	WLHSNVLARI DAAAFGLA	LEQLDLSDNA QLRVVDPTTF	HGLGRHLTLH
Murine NOGO-R3	WIYSNNITFI APNTFEGFVH	LEELDLDGDR QLRILAPETF	QGLVXLHALY
Human NOGO-R3	WIYSNNITFI HPSTFEGFVH	LEELDLDGDR QLRILAPETF	QGLVXLHALY
Human NOGO-R2	WLFSNNLSTI YPGTFRHLQ	LEELDLDGDR HLRSLPDTF	QGLERLQSLH
Consensus	W--SN--I	-----F-----	LE-LDL-DN--L--P-TF--GL--L--L--
	151	LRR 4	LRR 5
Human NOGO-R1	LDRCGQLQELG PGLFRGLAAL	QYLYLDNAL QALPDDTFRD	LGNLTHFLH
Murine NOGO-R1	LDRCGRLRELG PGLFRGLAAL	QYLYLDNHL QALPNTFRD	LGNLTHFLH
Murine NOGO-R3	LYKCGLSALP AGIFGGLHSL	QYLYLDNHI EYLQDDIFVD	LVNLSHLFLH
Human NOGO-R3	LYKCGLSALP AGVFGGLHSL	QYLYLDNHI EYLQDDIFVD	LVNLSHLFLH
Human NOGO-R2	LYRCQLSSLP GNIFRGLVSL	QYLYLDNSL LHLQDDLFAD	LVNLSHLFLH
Consensus	L--C-L--L--	---F-GL--L	QYLYLD--N--L--D--F-D-L--HLFLH

(57) Abstract: The invention relates generally to genes that encode proteins that inhibit axonal growth. The invention relates specifically to genes encoding NgR protein homologs in humans and mice. The invention also includes compositions and methods for modulating the expression and activity of Nogo and the NgR proteins. Specifically, the invention includes peptides, proteins and antibodies that block Nogo-mediated inhibition of axonal extension. The compositions and methods of the invention are useful in the treatment of cranial or cerebral trauma, spinal cord injury, stroke or a demyelinating disease.



IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(48) Date of publication of this corrected version:

15 May 2003

Published:

— *with international search report*

(15) Information about Correction:

see PCT Gazette No. 20/2003 of 15 May 2003, Section II

(88) Date of publication of the international search report:

23 January 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOGO RECEPTOR HOMOLOGS

FIELD OF THE INVENTION

5 The invention relates to neurology and molecular biology. More particularly, the invention relates to CNS neurons and axonal growth

BACKGROUND

10 Among the mechanisms through which the cells of an organism communicate with each other and obtain information and stimuli from their environment is through cell membrane receptor molecules expressed on the cell surface. Many such receptors have been identified, characterized, and sometimes classified into major receptor superfamilies based on structural motifs and signal transduction features. The receptors are a first essential link for translating an extracellular signal into a cellular
15 physiological response.

 Receptors on neurons are particularly important in the development of the nervous system during embryogenesis. The neurons form connections with target cells during development through axonal extension of the neurons toward the target cells in a receptor-mediated process. Axons and dendrites have a specialized region of their
20 distal tips known as the growth cone. Growth cones enable the neuron to sense the local environment through a receptor-mediated process and direct the movement of the axon or dendrite of the neuron toward the neuron's target cell. This process is known as elongation. Growth cones can be sensitive to several guidance cues, for example, surface adhesiveness, growth factors, neurotransmitters and electric fields. The
25 guidance of growth at the cone depends on various classes of adhesion molecules, intercellular signals, as well as factors that stimulate and inhibit growth cones.

 Interestingly, damaged neurons do not elongate in the central nervous system (CNS) following injury due to trauma or disease, whereas axons in the peripheral nervous system (PNS) regenerate readily. The fact that damaged CNS neurons fail to
30 elongate is not due to an intrinsic property of CNS axons, but rather due to the CNS environment that is not permissive for axonal elongation. Classical grafting experiments by Aguayo and colleagues (*e.g.*, Richardson *et al.*, (1980) *Nature* 284,

- 2 -

264-265) demonstrated that CNS axons can in fact elongate over substantial distances within peripheral nerve grafts, and that CNS myelin inhibits CNS axon elongation. Therefore, given the appropriate environment, CNS axons can regenerate, implying that CNS axonal injury can potentially be addressed by appropriate manipulation of the
5 CNS environment.

The absence of axon regeneration following injury can be attributed to the presence of axon growth inhibitors. These inhibitors are predominantly associated with myelin and constitute an important barrier to regeneration. Axon growth inhibitors are present in CNS-derived myelin and the plasma membrane of
10 oligodendrocytes that synthesize myelin in the CNS (Schwab *et al.*, (1993) *Annu. Rev. Neurosci.* 16, 565-595). Myelin-associated inhibitors appear to be a primary contributor to the failure of CNS axon regeneration *in vivo* after an interruption of axonal continuity, whereas other non-myelin associated axon growth inhibitors in the CNS may play a lesser role. These inhibitors block axonal regeneration following
15 neuronal injury due to trauma, stroke or viral infection.

Numerous myelin-derived axon growth inhibitors have been characterized (see, for review, David *et al.*, (1999) WO995394547; Bandman *et al.*, (1999) U.S. Patent No. 5,858,708; Schwab, (1996) *Neurochem. Res.* 21, 755-761). Several components of CNS white matter, NI35, NI250 (Nogo) and Myelin-associated glycoprotein
20 (MAG), which have inhibitory activity for axonal extension, have been described as well (Schwab *et al.*, (1990) WO9005191; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). In particular, Nogo is a 250 kDa myelin-associated axon growth inhibitor that was originally characterized based on the effects of the purified protein *in vitro* and monoclonal antibodies that neutralize the protein's activity (Schwab (1990) *Exp.*
25 *Neurol.* 109, 2-5). The Nogo cDNA was first identified through random analysis of brain cDNA and had no suggested function (Nagase *et al.*, (1998) *DNA Res.* 5, 355-364). The identification of this Nogo cDNA as the cDNA encoding the 250 kDa myelin-associated axon growth inhibitor was discovered only recently (GrandPre *et al.*, (2000) *Nature* 403, 439-444; Chen *et al.*, (2000) *Nature* 403, 434-439; Prinjha *et al.*,
30 (2000) *Nature* 403, 383-384).

Importantly, Nogo has been shown to be the primary component of CNS myelin responsible for inhibiting axonal elongation and regeneration. Nogo's selective

- 3 -

expression by oligodendrocytes and not by Schwann cells (the cells that myelinate P.S. axons) is consistent with the inhibitory effects of CNS myelin, in contrast to P.S. myelin (GrandPre *et al.*, (2000) *Nature* 403, 434-439). In culture, Nogo inhibits axonal elongation and causes growth cone collapse (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272, 19283-19293). Antibodies (*e.g.*, IN-1) against Nogo have been shown to block most of the inhibitory action of CNS myelin on neurite growth *in vitro* (Spillmann *et al.*, (1998) *J. Biol. Chem.* 272:19283-19293). These experiments indicate that Nogo is the main component of CNS myelin responsible for inhibition of axonal elongation in culture. Furthermore, *in vivo*, the IN-1 antibody has been shown to enhance axonal regeneration after spinal cord injury, resulting in recovery of behaviors such as contact placing and stride length (Schnell and Schwab (1990) *Nature* 343, 269-272; Bregman *et al.*, (1995) *Nature* 378, 498-501). Thus, there is substantial evidence that Nogo is a disease-relevant molecular target. Agents that interfere with the binding of Nogo to its receptor would be expected to improve axonal regeneration in clinical states in which axons have been damaged, and improve patient outcome.

Modulation of Nogo has been described as a means for treatment of regeneration for neurons damaged by trauma, infarction and degenerative disorders of the CNS (Schwab *et al.*, (1994) WO9417831; Tatagiba *et al.*, (1997) *Neurosurgery* 40, 541-546) as well as malignant tumors in the CNS such as glioblastoma (Schwab *et al.*, (1993) U.S. Patent No. 5,250,414); Schwab *et al.*, (2000) U.S. Patent No. 6,025,333).

Antibodies which recognize Nogo have been suggested to be useful in the diagnosis and treatment of nerve damage resulting from trauma, infarction and degenerative disorders of the CNS (Schnell & Schwab, (1990) *Nature* 343, 269-272; Schwab *et al.*, (1997) U.S. Patent No. 5,684,133). For CNS axons, there is a correlation between the presence of myelin and the inhibition of axon regeneration over long distances (Savio and Schwab (1990) *Proc. Natl. Acad. Sci.* 87, 4130-4133; Keirstead *et al.*, (1992) *Proc. Natl. Acad. Sci.* 89, 11664-11668). After Nogo is blocked by antibodies, neurons can again extend across lesions caused by nerve damage (Schnell and Schwab (1990) *Nature* 343, 269-272).

- 4 -

SUMMARY OF THE INVENTION

Genes encoding homologs (NgR2 and NgR3) of a Nogo receptor (NgR1) in mice and humans have been discovered. Various domains in the polypeptides encoded by the NgR2 and NgR3 genes have been identified and compared to domains in mouse and human NgR1 polypeptides. This comparison has led to identification of a consensus sequence (NgR consensus sequence) that characterizes a family of proteins (NgR family). Based on these and other discoveries, the invention features molecules and methods for modulating axonal growth in CNS neurons.

The invention provides a polypeptide that contains a polypeptide containing a tryptophan rich LRRCT domain consisting of the amino acid sequence:

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any protein amino acid or a gap, and the polypeptide does not include amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

Preferably, X17 and X23 are (independently) arginine or lysine. In some embodiments, the amino acid sequence of the LRRCT domain is residues 261-310 of SEQ ID NO:2, or residues 261-310 of SEQ ID NO: 2 with up to 10 conservative amino acid substitutions. In some embodiments, the polypeptide contains the following NTLRRCT amino acid sequence:

C P X₁ X₂ C X₃ C Y X₄ X₅ P X₆ X₇ T X₈ S C X₉ X₁₀ X₁₁ X₁₂ X₁₃ X₁₄ X₁₅ X₁₆ P
X₁₇ X₁₈ X₁₉ P X₂₀ X₂₁ X₂₂ X₂₃ R X₂₄ F L X₂₅ X₂₆ N X₂₇ I X₂₈ X₂₉ X₃₀ X₃₁ X₃₂ X₃₃
X₃₄ F X₃₅ X₃₆ X₃₇ X₃₈ X₃₉ X₄₀ X₄₁ X₄₂ L W X₄₃ X₄₄ S N X₄₅ X₄₆ X₄₇ X₄₈ I X₄₉
X₅₀ X₅₁ X₅₂ F X₅₃ X₅₄ X₅₅ X₅₆ X₅₇ L E X₅₈ L D L X₅₉ D N X₆₀ X₆₁ L X₆₂ X₆₃ X₆₄
X₆₅ P X₆₆ T F X₆₇ G L X₆₈ X₆₉ L X₇₀ X₇₁ L X₇₂ L X₇₃ X₇₄ C X₇₅ L X₇₆ X₇₇ L X₇₈
X₇₉ X₈₀ X₈₁ F X₈₂ G L X₈₃ X₈₄ L Q Y L Y L Q X₈₅ N X₈₆ X₈₇ X₈₈ X₈₉ L X₉₀ D

- 5 -

X₉₁ X₉₂ F X₉₃ D L X₉₄ N L X₉₅ H L F L H G N X₉₆ X₉₇ X₉₈ X₉₉ X₁₀₀ X₁₀₁ X₁₀₂
 X₁₀₃ X₁₀₄ F R G L X₁₀₅ X₁₀₆ L D R L L L H X₁₀₇ N X₁₀₈ X₁₀₉ X₁₁₀ X₁₁₁ V H X₁₁₂
 X₁₁₃ A F X₁₁₄ X₁₁₅ L X₁₁₆ R L X₁₁₇ X₁₁₈ L X₁₁₉ L F X₁₂₀ N X₁₂₁ L X₁₂₂ X₁₂₃ L
 X₁₂₄ X₁₂₅ X₁₂₆ X₁₂₇ L X₁₂₈ X₁₂₉ L X₁₃₀ X₁₃₁ L X₁₃₂ X₁₃₃ L R L N X₁₃₄ N X₁₃₅ W
 5 X₁₃₆ C X₁₃₇ C R X₁₃₈ R X₁₃₉ L W X₁₄₀ W X₁₄₁ X₁₄₂ X₁₄₃ X₁₄₄ R X₁₄₅ S S S X₁₄₆
 V X₁₄₇ C X₁₄₈ X₁₄₉ P X₁₅₀ X₁₅₁ X₁₅₂ X₁₅₃ X₁₅₄ X₁₅₅ D L X₁₅₆ X₁₅₇ L X₁₅₈ X₁₅₉ X₁₆₀
 D X₁₆₁ X₁₆₂ X₁₆₃ C [SEQ ID NO:18]

wherein X is any amino acid residue or a gap and wherein the polypeptide is not the
 10 polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1). For
 example, X₆, X₃₇ and X₃₈ may represent a gap. Specific examples of polypeptides of
 the invention are SEQ ID NO: 2 (human NgR2), SEQ ID NO: 4 (mouse NgR3), and
 SEQ ID NO: 14 (human NgR3). In some embodiments, the polypeptide contains: (a)
 a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial
 15 CTS domain, if present, consists of no more than the first 39 amino acids of the CTS
 domain. While the polypeptide may contain a functional GPI domain, a functional GPI
 domain may be absent, e.g., when a soluble polypeptide is desired. A polypeptide of
 the invention optionally includes an amino acid sequence of a heterologous
 polypeptide, e.g., an Fc portion of an antibody.

20 The invention also provides a nucleic acid encoding an above-described
 polypeptide; a vector containing the nucleic acid, which nucleic acid may be operably
 linked to an expression control sequence; and a transformed host cell containing the
 vector. A method of producing a polypeptide of the invention is also provided. The
 method includes introducing a nucleic acid encoding the above-described polypeptide
 25 into a host cell, culturing the cell under conditions suitable for expression of the
 polypeptide, and recovering the polypeptide.

The invention also provides an antisense molecule whose nucleotide sequence
 is complementary to a nucleotide sequence encoding a polypeptide selected from the
 group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a
 30 polypeptide consisting of residues 256-396 of SEQ ID NO:14 and a polypeptide
 consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to

- 6 -

100 nucleotides in length, e.g., about 20, 30, 40, 50, 60, 70, 80 or 90 nucleotides. The invention also provides a nucleic acid encoding such an antisense molecule.

The invention also provides an antibody that binds to an above-described polypeptide. Polypeptides or antibodies of the invention can be formulated into
5 pharmaceutical compositions containing the polypeptide or antibody and a pharmaceutically acceptable carrier.

The invention also provides a method for decreasing inhibition of axonal growth of a CNS neuron. The method includes the step of contacting the neuron with an effective amount of a polypeptide or antibody of the invention.

10 The invention also provides a method for treating a central nervous system disease, disorder or injury. The method includes administering to a mammal, e.g., a human, an effective amount of a polypeptide or antibody of the invention. Exemplary diseases, disorders and injuries that may be treated using molecules and methods of the invention include, but are not limited to, cerebral injury, spinal cord injury, stroke,
15 demyelinating diseases, e.g., multiple sclerosis, monophasic demyelination, encephalomyelitis, multifocal leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease, Spongy degeneration, Alexander's disease, Canavan's disease, metachromatic leukodystrophy and Krabbe's disease.

The invention also provides a method for identifying a molecule that binds a
20 polypeptide of the invention. The method includes the steps of: (a) providing a polypeptide of the invention; (b) contacting the polypeptide with the candidate molecule; and (c) detecting binding of the candidate molecule to the polypeptide.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the
25 invention belongs. In case of conflict, the present application, including definitions, will control. All publications, patent and other references mentioned herein are incorporated by reference.

The materials, methods and examples presented below are illustrative only, and not intended to be limiting. Other features and advantages of the invention will be
30 apparent from the detail description and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

- 7 -

Fig. 1A-1B shows an alignment of NgR2 (SEQ ID NO:2) and NgR3 (SEQ ID NO:4) with the known NgR, NgR1 (SEQ ID NO:5) and the Consensus Sequence (SEQ ID NO:6).

5

Fig. 2. mNgR3 does not bind hNogoA(1055-1120). COS-7 cells were transfected with vectors encoding myc-NgR1 or myc-NgR3, fixed, and stained with anti-myc antibodies or AP-hNogoA(1055-1120).

10

Fig.3. An alignment of the amino acid sequences of human NgR1, murine NgR1, murine NgR3, human NgR3 and human NgR2. Numbering begins with amino acid #1 of murine NgR3. The consensus sequence is listed below. The LLR NT domain is indicated by a shaded box; domains LLR 1, LLR 3, LLR 5, and LLR 7 are indicated by open boxes; LLR 2, LLR 4, LLR 6 and LLR 8 are indicated by shaded boxes; and the LLR CT domain is indicated by a shaded box. Amino acids in bold in LLR 8 indicate a conserved glycosylation sites. A dot indicates conserved cystine residue in LLR4. Box at C terminus indicates putative GPI signals.

15

20

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) encoding NgR homologs, referred to herein as NgR. Unless indicated otherwise, as used herein, the abbreviation in lower case (NgR) refers to a gene, cDNA, RNA or nucleic acid sequence, whereas the upper case version (NgR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence. Specific proteins are designated by number, *e.g.*, "NgR2" is a human NgR homolog, "NgR3" is a murine-derived NgR homolog, and "NgR1" is the known NgR identified by Dr. Stephen Strittmatter. Known NgRs are herein referred to as "NgRs." DNA polynucleotides of the invention include genomic DNA, cDNA and DNA that has been chemically synthesized in whole or in part.

25

30

- 8 -

Standard reference works setting forth the general principles of recombinant DNA technology known to those of skill in the art include Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York (1998); Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d Ed., Cold Spring Harbor Laboratory Press, Plainview, New York (1989); Kaufman *et al.*, Eds., HANDBOOK OF MOLECULAR AND CELLULAR METHODS IN BIOLOGY AND MEDICINE, CRC Press, Boca Raton (1995); McPherson, Ed., DIRECTED MUTAGENESIS: A PRACTICAL APPROACH, IRL Press, Oxford (1991).

As used herein, the term "axon" refers to a long cellular protrusion from a neuron, whereby action potentials are conducted, either to or from the cell body.

As used herein, the term "axonal growth" refers to an extension of the long process or axon, originating at the cell body and proceeded by the growth cone.

As used herein, the term "central nervous system disorder" refers to any pathological state associated with abnormal function of the central nervous system (CNS). The term includes, but is not limited to, altered CNS function resulting from physical trauma to cerebral tissue, viral infection, autoimmune mechanisms and genetic mutation.

As used herein, the term "demyelinating disease" refers to a pathological disorder characterized by the degradation of the myelin sheath of the oligodendrocyte cell membrane.

As used herein, the term "growth cone" refers to a specialized region at the tip of a growing neurite that is responsible for sensing the local environment and moving the axon toward its appropriate synaptic target cell.

As used herein, the term "growth cone movement" refers to the extension or collapse of the growth cone toward a neuron's target cell.

As used herein, the term "neurite" refers to a process growing out of a neuron. As it is sometimes difficult to distinguish a dendrite from an axon in culture, the term "neurite" is used for both.

As used herein, the term "oligodendrocyte" refers to a neuroglial cell of the CNS whose function is to myelinate CNS axons.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods.

- 9 -

"Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means. By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein. The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region. Examples of NgR protein domains include, but are not limited to, the signal peptide, extracellular (*i.e.*, N-terminal) domain, and leucine-rich repeat domains.

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, *i.e.*, having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event. Such activities may be measured by assays such as competitive inhibition of NgR1 binding to Nogo assays wherein, for example, unlabeled, soluble NgR2 is added to an assay system in increasing concentrations to inhibit the binding of Nogo to NgR1 expressed on the surface of CHO cells. As another example, one may assess the ability of neurons to extend across lesions caused by nerve damage (as in Schnell and Schwab (1990) *Nature* 343, 269-272) following inhibition of Nogo by various forms of NgR2 and/or NgR3 as a biological indicator of NgR function.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies, anti-idiotypic antibodies, anti-anti-idiotypic antibodies, and humanized antibodies.

As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or

- 10 -

combinations thereof. Binding includes ionic, non-ionic, hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or
5 due to the effect of another protein or compound but instead are without other substantial chemical intermediates.

As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecules, peptides, proteins, sugars, nucleotides or nucleic acids, and such compound can be natural or synthetic.

10 As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers,
15 salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the NgR or fragment thereof.

As used herein, the phrase "homologous nucleotide sequence," or "homologous
20 amino acid sequence," or variations thereof, refers to sequences characterized by an identity at the nucleotide level, or a homology at the amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA.

25 Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does
30 not, however, include the nucleotide sequence encoding NgR1. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. A

- 11 -

homologous amino acid sequence does not, however, include the amino acid sequence encoding other known NgRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that is substantially free of nucleic acids encoding other proteins with which it is associated in nature, i.e., a nucleic acid that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NgR nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

As used herein, the term "heterologous" refers to a nucleotide or amino acid sequence that is a different, or non-corresponding sequence, or a sequence derived from a different species. For example, a mouse NgR nucleotide or amino acid sequence is heterologous to a human NgR nucleotide or amino acid sequence, and a human NgR nucleic or amino acid sequence is heterologous to a human immunoglobulin nucleotide or amino acid sequence.

As used herein, a "soluble NgR polypeptide" is a NgR polypeptide that does not anchor itself in a membrane. Such soluble polypeptides include, for example, NgR2 and NgR3 polypeptides that lack a sufficient portion of their GPI anchor signal

- 12 -

to anchor the polypeptide or are modified such that the GPI anchor signal is not adequate to result in replacement of the peptide with a GPI anchor. In preferred embodiments, up to 5, 10, 20 or 25 amino acids are removed from the C-terminus of NgR2 or NgR3 to make the respective proteins soluble. As used herein soluble NgR polypeptides include full-length or truncated (*e.g.*, with internal deletions) NgR.

Soluble NgR polypeptides may include the entire NgR protein up to the putative GPI signal sequence (*e.g.*, amino acid 1 to about amino acid 395 of NgR2, and from amino acid 1 to about amino acid 438 of NgR3). In other embodiments, the signal peptide of the proteins may be removed or truncated (*e.g.*, all or part of the signal sequence of NgR2, which spans amino acid 1 to about amino acid 30 of SEQ ID NO:2, may be removed; all or part of the signal sequence of NgR3, which spans amino acid 1 to about amino acid 40 of SEQ ID NO:4, may be removed). In some embodiments, the mature NgR2 (SEQ ID NO:8) and the mature NgR3 (SEQ ID NO:9) are used.

Soluble NgR polypeptides include at least one of the putative ligand-binding portions of NgR, including the first cysteine-rich region (SEQ ID NO:10, the leucine repeat region (SEQ ID NO:12) and the second cysteine-rich region (SEQ ID NO:11). In some embodiments, soluble NgR polypeptides consist of amino acid 1 through about amino acid 395 of SEQ ID NO:2, or amino acid 1 through about amino acid 438 of SEQ ID NO:4.

In other embodiments, the soluble NgR polypeptides are fusion proteins that contain amino acids 30 through about amino acid 395 of mature NgR2 or amino acid 40 through about amino acid 438 of NgR3, the C-terminal 10 amino acids of a human IgG 1 hinge region containing the two cysteine residues thought to participate in interchain disulfide bonding, and the CH2 and CH3 regions of a human IgG1 heavy chain constant domain. This type of recombinant protein is designed to modulate inhibition of axonal elongation through inhibition of the Nogo ligand binding to NgR1, or by inhibiting the ligand of the NgR from interacting with cell surface NgR. The NgR portion of the fusion binds to the Nogo ligand and the IgG1 portion binds to the FcγRI (macrophage) and FcγIII (NK cells and neutrophils) receptors.

The production of the soluble polypeptides useful in this invention may be achieved by a variety of methods known in the art. For example, the polypeptides may

- 13 -

be derived from intact transmembrane NgR molecules by proteolysis using specific endopeptidases in combination with exopeptidases, Edman degradation, or both. The intact NgR molecule, in turn, may be purified from its natural source using conventional methods. Alternatively, the intact NgR may be produced by known
5 recombinant DNA techniques using cDNAs, expression vectors and well-known techniques for recombinant gene expression.

Preferably, the soluble polypeptides useful in the present invention are produced directly, thus eliminating the need for an entire NgR as a starting material. This may be achieved by conventional chemical synthesis techniques or by well-known
10 recombinant DNA techniques wherein only those DNA sequences which encode the desired peptides are expressed in transformed hosts. For example, a gene which encodes the desired soluble NgR polypeptide may be synthesized by chemical means using an oligonucleotide synthesizer. Such oligonucleotides are designed based on the amino acid sequence of the desired soluble NgR polypeptide. Specific DNA sequences
15 coding for the desired peptide also can be derived from the full-length DNA sequence by isolation of specific restriction endonuclease fragments or by PCR synthesis of the specified region from cDNA.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NOs:1, 3 or a complement of either of
20 these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NOs:1 or 3 as a hybridization probe, NgR nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY
25 MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993).

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers
30 according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

- 14 -

Furthermore, oligonucleotides corresponding to NgR nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

5 As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue.

10 Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

 As used herein, the term "probe" refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides,
15 depending on use. They are used in the detection of identical, similar or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

20 The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

 The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

 The term "therapeutic effect" refers to the inhibition or activation factors
25 causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of
30 degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population

- 15 -

of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions, for example, include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus and inflammation.

Abnormal differentiation conditions include, for example, neurodegenerative disorders, slow wound healing rates and slow tissue grafting healing rates.

Abnormal cell signaling conditions include, for example, psychiatric disorders involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then

- 16 -

be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable
5 presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1–2-fold, and preferably more, compared to the basal level.

The amino acid sequences are presented in the amino to carboxy direction,
10 from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

15

Nucleic Acids

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA
20 transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a NgR polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the
25 invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild-type genes, are naturally occurring sequences (as
30 opposed to non-naturally occurring variants arising from in vitro manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding NgR (conventionally followed by

- 17 -

second-strand synthesis of a complementary strand to provide a double-stranded DNA).

Preferred DNA sequences encoding a human NgR polypeptide is set out in SEQ ID NOs:1 and 13. A preferred DNA of the invention comprises a double stranded molecule comprising the coding molecule (*i.e.*, the "coding strand") along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding NgR polypeptides, as shown in SEQ ID NO:3, which comprises murine NgR homolog, NgR3.

Also preferred are nucleotide sequences that encode at least a portion of a NgR polypeptide that has at least one biological function of a NgR. More preferred are nucleotide sequences that encode a portion of NgR that encodes at least the mature NgR without the hydrophobic C-terminal GPI signal. Also preferred are nucleotide sequences that encode the portion of NgR that encodes at least the ligand-binding region of NgR.

The invention further embraces other species, preferably mammalian, homologs of the human NgR DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with human DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the NgR sequences set forth in SEQ ID NOs:1, 3 or 13, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art. Polynucleotides of the invention also permit identification and isolation of polynucleotides encoding related NgR polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR).

- 18 -

Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to NgR and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of NgR. Non-human species genes encoding proteins homologous to NgR can also be identified by Southern and/or PCR analysis and are useful in animal models for NgR disorders. Knowledge of the sequence of a human NgR DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding NgR expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express NgR. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a NgR locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

The disclosure herein of a full-length polynucleotide encoding a NgR polypeptide makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide. The invention, therefore, provides fragments of NgR-encoding polynucleotides comprising at least 6, and preferably at least 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding NgR. Preferably, fragments of polynucleotides of the invention comprise sequences unique to the NgR-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding NgR (or fragments thereof).

Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic

- 19 -

DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent and enzymatic labeling.

5 Fragments of polynucleotides are particularly useful as probes for detection of full-length or fragment of NgR polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding NgR, or used to detect variations in a polynucleotide sequence encoding NgR.

10 The invention also embraces DNAs encoding NgR polypeptides that hybridize under moderately stringent or high stringency conditions to the noncoding strand, or complement, of the polynucleotide in any of SEQ ID NOs:1 or 3.

15 Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1/6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98% or 99% homologous to each other typically remain
20 hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that
25 hybridizes under stringent conditions to the sequence of SEQ ID NOs:1 or 3 corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein). As used herein, "stringent hybridization conditions" means: 42°C in a hybridization
30 solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% (wt/vol) dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS.

Vectors

30 Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding NgR and/or to

- 20 -

express DNA which encodes NgR. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses); that serve equivalent functions.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67, 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia,

- 21 -

Piscataway, N.J.) that fuse glutathione-S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69, 301-315) and pET 11d (Studier *et al.*, GENE
5 EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY:
10 METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20, 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by
15 standard DNA synthesis techniques.

In another embodiment, the NgR expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J.* 6, 229-234), pMFa (Kurjan and Herskowitz (1982) *Cell* 30, 933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54, 113-123), pYES2
20 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, NgR can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith *et al.*, (1983) *Mol. Cell. Biol.* 3,
25 2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170, 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329, 840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6, 187-195). When used in mammalian cells, the
30 expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both

- 22 -

prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, (Eds.) MOLECULAR CLONING: A LABORATORY MANUAL. 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is
5 capable of directing expression of the nucleic acid preferentially in a particular cell type
(*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-
specific regulatory elements are known in the art. Non-limiting examples of suitable
tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*
(1987) *Genes Dev.* 1, 268-277), lymphoid-specific promoters (Calame and Eaton
10 (1988) *Adv. Immunol.* 43, 235-275), in particular promoters of T cell receptors
(Winoto and Baltimore (1989) *EMBO J.* 8, 729-733) and immunoglobulins (Banerji *et al.*
(1983) *Cell* 33, 729-740; Queen and Baltimore (1983) *Cell* 33, 741-748), neuron-
specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc.*
Natl. Acad. Sci. USA 86, 5473-5477), pancreas-specific promoters (Edlund *et al.*
15 (1985) *Science* 230, 912-916), and mammary gland-specific promoters (*e.g.*, milk
whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No.
264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the
murine hox promoters (Kessel and Gruss (1990) *Science* 249, 374-379) and the
 α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3, 537-546).

20 The invention further provides a recombinant expression vector comprising a
DNA molecule of the invention cloned into the expression vector in an antisense
orientation. That is, the DNA molecule is operatively linked to a regulatory sequence
in a manner that allows for expression (by transcription of the DNA molecule) of an
RNA molecule that is antisense NgR mRNA. Regulatory sequences operatively linked
25 to a nucleic acid cloned in the antisense orientation can be chosen that direct the
continuous expression of the antisense RNA molecule in a variety of cell types, for
instance viral promoters and/or enhancers, or regulatory sequences can be chosen that
direct constitutive, tissue-specific or cell-type-specific expression of antisense RNA.
The antisense expression vector can be in the form of a recombinant plasmid,
30 phagemid or attenuated virus in which antisense nucleic acids are produced under the
control of a high efficiency regulatory region, the activity of which can be determined
by the cell type into which the vector is introduced. For a discussion of the regulation

- 23 -

of gene expression using antisense genes see Weintraub *et al.*, *Antisense RNA as a molecular tool for genetic analysis*, REVIEWS--TRENDS IN GENETICS, Vol. 1(1) 1986.

Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses and integratable DNA fragments (*i.e.*, fragments
5 integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors
10 include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen) and pYES2™ (Invitrogen).

Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding NgR is operably linked or connected to suitable control sequences
15 capable of effecting the expression of the NgR in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually
20 conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include, but are not limited to a transcriptional promoter, enhancers, an optional operator sequence to control transcription, polyadenylation signals, a
25 sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation. Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types
30 of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as

- 24 -

the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., NgR proteins, mutant forms of NgR, fusion proteins, etc.).

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the PR and PL promoters of bacteriophage lambda (THE BACTERIOPHAGE LAMBDA, Hershey, A.D. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; LAMBDA II, Hendrix, R.W. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the trp, recA, heat shock, and lacZ promoters of *E. coli* and the SV40 early promoter (Benoist *et al.*, (1981) *Nature* 290, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine and human metallothionein.

Additional regulatory sequences can also be included in preferred vectors. Preferred examples of suitable regulatory sequences are represented by the Shine-Dalgarno sequence of the replicase gene of the phage MS-2 and of the gene cII of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding NgR and result in the expression of the mature NgR protein.

Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors

- 25 -

which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and NgR DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

5 Nucleotide sequences encoding NgR may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook *et al.*, *supra* and are well known in the art. Methods for
10 construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, (1983) *Mol. Cell. Biol.* 3:280, Cosman *et al.* (1986) *Mol. Immunol.* 23:935, Cosman *et al.*, (1984) *Nature* 312:768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

15

Host Cells and Transformed Host Cells

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded NgR
20 polypeptide. Preferably, the cell produces little or no endogenous NgR polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation,
25 nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate and mammalian cells systems.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with NgR. Host cells of the
30 invention are also useful in methods for the large-scale production of NgR polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells

- 26 -

are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of NgR DNA sequences allows for modification of cells to permit, or increase, expression of endogenous NgR. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring NgR promoter with all or part of a heterologous promoter so that the cells express NgR at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous NgR encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the NgR coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the NgR coding sequences in the cells.

The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination or "knock-out" strategies; see Capecchi, *Science* 244:1288-1292 (1989)) of animals that fail to express functional NgR or that express a variant of NgR. Such animals (especially small laboratory animals such as rats, rabbits and mice) are useful as models for studying the *in vivo* activities of NgR and modulators of NgR.

- 27 -

Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces* and *Staphylococcus*.

If a eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*, Tissue Culture, Academic Press, Kruse and Patterson, Eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast cell may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia* and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, 1988, 6, 47; BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, 1992; and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, CA

- 28 -

(1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via
5 conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or
10 transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells
15 may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin, dihydrofolate reductase (DHFR) and methotrexate. Nucleic acid
20 encoding a selectable marker can be introduced into a host cell on the same vector as that encoding NgR or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

In a preferred embodiment, the polypeptides of the invention, including forms
25 of NgR2 and NgR3, soluble forms of NgR, chimeric NgR polypeptides, NgR/Ig fusions and fragments and variations of each of the above are expressed in Chinese Hamster Ovary (CHO) cells.

In order to introduce the DNA fragment coding for the NgR protein or polypeptide into the CHO cell to express the recombinant NgR protein or polypeptide,
30 it is necessary to construct the expression vector.

The vectors for CHO expression include, but are not limited to, pA1-11, pXT1, pRc/CMV, pRc/RSV and pcDNAIneo. The promoter is not specifically limited

- 29 -

provided it effectively promotes expression in CHO cells. Examples of suitable promoters are: SR α , SV40, LTR, CMV, and HSV-TK. Of these, CMV and Sr α promoters are preferred.

5 In addition to the above-mentioned promoters, the expression vectors may contain enhancers, splicing signals, polyadenylation signals, selectable markers and an SV40 replication origin. Suitable selectable markers include, but are not limited to the dihydrofolate reductase (DHFR) gene which provides resistance to methotrexate (MTX), the ampicillin resistance gene, and the neomycin resistance gene.

10 Examples of the expression vectors each containing the DNA coding for NgR, portions, fragments and soluble constructs thereof, include the vector (such as one described above), into which the promoter is operably linked (preferably upstream) to the nucleotide sequence encoding the desired NgR construct; a polyadenylation signal downstream from the nucleotide sequence encoding the NgR construct; and, preferably, the vector includes an operable DHFR gene. Preferably, the ampicillin
15 resistant gene is also operably contained in the vector.

CHO cell lacking the DHFR gene (Urlaub, G. *et al.*, (1980) *Proc. Natl. Acad. Sci. USA* 77, 4216-4220) and CHO-K1 (*Proc. Natl. Acad. Sci. USA* 60, 1275 (1968)) are suitable for use.

20 The NgR expression vectors prepared as above are introduced into CHO cells by any known method, including, but not limited to the calcium phosphate method (Graham and van der Eb (1973) *Virol.* 52, 456-467) and electroporation (Nuemann *et al.*, (1982) *EMBO J.* 1, 841-845).

Transformants carrying the expression vectors are selected based on the above-mentioned selectable markers. Repeated clonal selection of the transformants
25 using the selectable markers allows selection of stable cell lines having high expression of the NgR constructs. Increased MTX concentrations in the selection medium allows gene amplification and greater expression of the desired protein. The CHO cell containing the recombinant NgR can be produced by cultivating the CHO cells containing the NR expression vectors constitutively expressing the NgR constructs.

30 Media used in cultivating CHO cells includes DMEM medium supplemented with about 0.5 to 20% fetal calf serum, DMEM medium and RPMI1640 medium. The

- 30 -

pH of the medium is preferably about 6 to 8. Cultivation is preferably at about 30 to 40°C for about 15 to 72 hours with aeration.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NgR protein. Accordingly, the invention further provides methods for producing NgR protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding NgR has been introduced) in a suitable medium such that NgR protein is produced. In another embodiment, the method further comprises isolating NgR from the medium or the host cell.

In situations where the NgR polypeptide will be found primarily intracellularly, intracellular material (including inclusion bodies for Gram-negative bacteria) can be extracted from the host cell using any standard technique known to one of ordinary skill in the art. Such methods would encompass, by way of example and not by way of limitation, lysing the host cells to release the contents of the periplasm/cytoplasm by French press, homogenization and/or sonication followed by centrifugation.

If the NgR polypeptide has formed inclusion bodies in the cytosol, such inclusion bodies may frequently bind to the inner and/or outer cellular membranes. Upon centrifugation, the inclusion bodies will be found primarily in the pellet material. The pellet material can then be treated at pH extremes or with one or more chaotropic agents such as a detergent, guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris-carboxyethyl phosphine at acid pH to release, break apart and solubilize the inclusion bodies. Once solubilized, NgR polypeptide can be analyzed using gel electrophoresis, immunoprecipitation or the like. Various methods of isolating the NgR polypeptide would be apparent to one of ordinary skill in the art, for example, isolation may be accomplished using standard methods such as those set forth below and in Marston *et al* (1990) *Meth. Enzymol.* 182, 264-275 (incorporated by reference herein in its entirety).

If isolated NgR polypeptide is not biologically active following the isolation procedure employed, various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages, can be used to restore biological

- 31 -

activity. Methods known to one of ordinary skill in the art include adjusting the pH of the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization but usually at a lower concentration and is not necessarily the same chaotrope as used for the solubilization. It may be required to employ a reducing agent or the reducing agent plus its oxidized form in a specific ratio, to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridge(s). Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol (DTT)/dithiane DTT, 2-mercaptoethanol (bME)/dithio-b(ME). To increase the efficiency of the refolding, it may be necessary to employ a cosolvent, such as glycerol, polyethylene glycol of various molecular weights and arginine.

Transgenic Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NgR-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NgR sequences have been introduced into their genome or homologous recombinant animals in which endogenous NgR sequences have been altered. Such animals are useful for studying the function and/or activity of NgR and for identifying and/or evaluating modulators of NgR activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous NgR gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA

- 32 -

molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NgR-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by
5 microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human NgR DNA sequence of SEQ ID NOs:1 or 3 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homolog of the human NgR gene, such as a mouse NgR gene, can be isolated based on hybridization to the human NgR cDNA (described
10 further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the NgR transgene to direct expression of NgR protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly
15 animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, in MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the
20 NgR transgene in its genome and/or expression of NgR mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding NgR can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which
25 contains at least a portion of a NgR gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NgR gene. The NgR gene can be a human gene (*e.g.*, SEQ ID NOs:1 or 13), but more preferably, is a non-human homolog of a human NgR gene. For example, a mouse homolog of human NgR gene of SEQ ID NOs:1 or 13 can be used to construct a homologous
30 recombination vector suitable for altering an endogenous NgR gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous

- 33 -

recombination, the endogenous NgR gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NgR gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NgR protein). In the homologous recombination vector, the altered portion of the NgR gene is flanked at its 5' and 3' ends by additional nucleic acid of the NgR gene to allow for homologous recombination to occur between the exogenous NgR gene carried by the vector and an endogenous NgR gene in an embryonic stem cell. The additional flanking NgR nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NgR gene has homologously recombined with the endogenous NgR gene are selected (see *e.g.*, Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. See *e.g.*, Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A Practical Approach, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr. Opin. Biotechnol.* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For

- 34 -

a description of the cre/loxP recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate
5 expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

10 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal
15 of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

20

Antisense

Also provided by the invention are antisense polynucleotides that recognize and hybridize to NgR polynucleotides. Full-length and fragment antisense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that
25 specifically recognize and hybridize to NgR RNA (as determined by sequence comparison of DNA encoding NgR to DNA encoding other known molecules). Identification of sequences unique to NgR encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the
30 desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be

- 35 -

performed. Antisense polynucleotides are particularly relevant to regulating expression of NgR by those cells expressing NgR mRNA.

Antisense oligonucleotides, or fragments of a nucleotide sequence set forth in SEQ ID NO:1, 3, 13 or sequences complementary or homologous thereto, derived
5 from the nucleotide sequences of the present invention encoding NgR are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. In specific aspects, antisense
10 nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire NgR coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a NgR protein of SEQ ID NO:2, 4 or 14 or antisense nucleic acids complementary to a NgR nucleic acid sequence of SEQ ID NOs:1, 3 or
15 13 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding NgR. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the protein coding region of human
20 NgR corresponds to the coding region SEQ ID NO:1, 3 or 13). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding NgR. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

25 Antisense oligonucleotides are preferably directed to regulatory regions of a nucleotide sequence of SEQ ID NO:1, 3, 13 or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like. Given the coding strand sequences encoding NgR disclosed herein (*e.g.*, SEQ ID NO:1, 3 or 13), antisense nucleic acids of the invention can be designed according
30 to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NgR mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding

- 36 -

or noncoding region of NgR mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NgR mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be
5 constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the
10 antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-
15 carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-
20 mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (*v*), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (*v*), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)*w*, and 2,6-
25 diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

30 The antisense nucleic acid molecules of the invention (preferably oligonucleotides of 10 to 20 nucleotides in length) are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA

- 37 -

and/or genomic DNA encoding a NgR protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. Suppression of NgR expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant NgR expression. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix.

Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine or cholesterol moieties at their 5' end.

An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.*, (1987) *Nucleic Acids Res.* 15, 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, (1987) *Nucleic Acids Res.* 15, 6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.*, (1987) *FEBS Lett.* 215, 327-330).

The NgR sequences taught in the present invention facilitate the design of novel transcription factors for modulating NgR expression in native cells and animals,

- 38 -

and cells transformed or transfected with NgR polynucleotides. For example, the Cys₂-His₂ zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular NgR target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal *et al.*, (1999) *Proc. Natl. Acad. Sci. USA* 96, 2758-2763; Liu *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 5525-5530; Greisman *et al.* (1997) *Science* 275, 657-661; Choo *et al.*, (1997) *J. Mol. Biol.* 273, 525-532). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to ensure specificity for a particular sequence (Segal *et al.*, (1999), above). The artificial zinc finger repeats, designed based on the promoter of NgR sequences, are fused to activation or repression domains to promote or suppress NgR expression (Liu *et al.*, (1997), above). The promoter of NgR may be obtained by standard methods known to one of ordinary skill in the art with the disclosure contained herein and knowledge of the NgR sequence. Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 3616-3620. Such proteins and polynucleotides that encode them, have utility for modulating NgR expression *in vivo* in both native cells, animals and humans; and/or cells transfected with NgR-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 96, 9521-9526; Wu *et al.*, (1995) *Proc. Natl. Acad. Sci. USA* 92, 344-348). The present invention contemplates methods of

- 39 -

designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate NgR expression in cells (native or transformed) whose genetic complement includes these sequences.

5 **Ribozymes and PNA moieties**

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes,
10 described in Haselhoff and Gerlach (1988) *Nature* 334, 585-591) can be used to catalytically cleave NgR mRNA transcripts to thereby inhibit translation of NgR mRNA. A ribozyme having specificity for a NgR-encoding nucleic acid can be designed based upon the nucleotide sequence of a NgR DNA disclosed herein (*i.e.*, SEQ ID NOs:1, 3 or 13). For example, a derivative of a Tetrahymena L-19 IVS RNA
15 can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a NgR-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, NgR mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*,
20 (1993) *Science* 261, 1411-1418.

Alternatively, NgR gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NgR (*e.g.*, the NgR promoter and/or enhancers) to form triple helical structures that prevent transcription of the NgR gene in target cells. See generally, Helene (1991) *Anticancer Drug Des.* 6:
25 569-584; Helene. *et al.*, (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *BioEssays* 14, 807-815.

In various embodiments, the nucleic acids of NgR can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate
30 backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.*, (1996) *Bioorg. Med. Chem. Lett.* 4, 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in

- 40 -

which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard
5 solid phase peptide synthesis protocols as described in Hyrup *et al.*, (1996) above; Perry-O'Keefe *et al.*, (1996) *Proc. Natl. Acad. Sci. USA* 93,14670-14675.

PNAs of NgR can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or
10 inhibiting replication. PNAs of NgR can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup (1996), above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.*, (1996), above; Perry-O'Keefe (1996), above).

15 In another embodiment, PNAs of NgR can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of NgR can be generated that may combine the advantageous properties of PNA and DNA. Such
20 chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), above). The synthesis of PNA-DNA chimeras can be
25 performed as described in Hyrup (1996), above and Finn *et al.* (1996) *Nucleic Acids Res.* 24, 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl) amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucleic Acids Res.* 17,
30 973-988). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996), above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment

- 41 -

and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg. Med. Chem. Lett.* 5:1119-1124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see Letsinger *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86, 6553-6556; Lemaitre *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (see, *e.g.*, Krol *et al.*, (1988) *Biotechniques* 6, 958-976) or intercalating agents (see, *e.g.*, Zon (1988) *Pharm. Res.* 5, 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

Automated sequencing methods can be used to obtain or verify the nucleotide sequence of NgR. The NgR nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

Polypeptides

The invention also provides purified and isolated mammalian NgR polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human NgR polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or SEQ ID NO:14. Another preferred embodiment is a mouse NgR polypeptide comprising the amino acid sequence of NgR3, as set forth in SEQ ID NO:4.

- 42 -

One aspect of the invention pertains to isolated NgR proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-NgR antibodies. Preferably, fragments of NgR proteins comprise at least one biological activity of NgR. In one embodiment, native NgR proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NgR proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a NgR protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, at least 50% or at least 45% identity and/or homology to the preferred polypeptide of the invention. In addition, the invention embraces polypeptides having the consensus sequence shown in SEQ ID NO:6, shown in Table 5) excluding the previously characterized NgR ("NgR1"), and polypeptides comprising at least about 90% of the consensus sequence.

The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment (Dayhoff, in ATLAS OF PROTEIN SEQUENCE AND

- 43 -

STRUCTURE, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference).

A determination of homology or identity is typically made by a computer homology program known in the art. An exemplary program is the Gap program
5 (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (*Adv. Appl. Math.*, 1981, 2, 482-489, which is incorporated herein by reference in its entirety). Employing the GAP software provided in the GCG program package, (see *Needleman and Wunsch* (1970) *J. Mol.*
10 *Biol.* 48, 443-453) the following settings for nucleic acid sequence comparison may be used: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NOs:1, 3 or 13.
15 BestFit was originally written for Version 1.0 by Paul Haeberli from a careful reading of the papers by Needleman and Wunsch (1970), above, and Smith and Waterman (1981), above. The following Bestfit settings for nucleic acid sequence comparison may be used: GAP creation penalty of 8.0 and GAP extension penalty of 2, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of
20 identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99%, with the CDS (encoding) part of the amino acid sequence shown in SEQ ID NOs:2, 4 or 14.

Alternatively, homology may be determined by hybridization analysis wherein a nucleic acid sequence is hybridized to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent
25 conditions. See *e.g.* Ausubel, *et al.*, (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention.

30 An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NgR protein is derived, or substantially free from

- 44 -

chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of NgR protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NgR protein having less than about 30% (by dry weight) of non-NgR protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-NgR protein, still more preferably less than about 10% of non-NgR protein, and most preferably less than about 5% non-NgR protein. When the NgR protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of NgR protein having less than about 30% (by dry weight) of chemical precursors or non-NgR chemicals, more preferably less than about 20% chemical precursors or non-NgR chemicals, still more preferably less than about 10% chemical precursors or non-NgR chemicals, and most preferably less than about 5% chemical precursors or non-NgR chemicals.

Biologically active portions of a NgR protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the NgR protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, 4 or 14 that include fewer amino acids than the full length NgR proteins, and exhibit at least one activity of a NgR protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the NgR protein. A biologically active portion of a NgR protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a NgR protein of the present invention may contain at least one of the features that is conserved between the NgR proteins (*e.g.*, a conserved cysteine as the N-terminus of the mature protein, four conserved cysteines

- 45 -

in the N-terminus before a leucine-rich region, four conserved cysteines C-terminal with respect to a leucine repeat region, eight leucine-rich repeats, and a hydrophobic C-terminus). An alternative biologically active portion of a NgR protein may contain at least two of the above-identified domains. Another biologically active portion of a
5 NgR protein may contain at least three of the above-identified domains. Yet another biologically active portion of a NgR protein of the present invention may contain at least four of the above-identified domains.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one
10 or more of the functional activities of a native NgR protein.

In an embodiment, the NgR protein has an amino acid sequence shown in SEQ ID NO:2, 4 or 14. In other embodiments, the NgR protein is substantially homologous to SEQ ID NO:2, 4 or 14 and retains the functional activity of the protein of SEQ ID NO:2, 4 or 14, yet differs in amino acid sequence due to natural allelic variation or
15 mutagenesis, as described in detail below.

Accordingly, in another embodiment, the NgR protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4 or SEQ ID NO:14 and retains the functional activity of the NgR proteins of SEQ ID NO:2, 4 or 14.

20 Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of NgR polypeptides are embraced by the invention.

25 The invention also embraces variant (or analog) NgR polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a NgR amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the NgR amino acid sequence. Insertional variants with additional residues at either or both termini
30 can include, for example, fusion proteins and proteins including amino acid tags or labels.

- 46 -

Insertion variants include NgR polypeptides wherein one or more amino acid residues are added to a NgR acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature NgR products, *i.e.*, NgR products wherein leader or signal sequences are removed, with additional amino
5 terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. NgR products with an additional methionine residue at position -1 (Met⁻¹-NgR) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met⁻²-Lys⁻¹-NgR). Variants of NgR with
10 additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

Polypeptide Variants

The invention also embraces NgR variants having additional amino acid
15 residues which result from use of specific expression systems.

As used herein, a NgR "chimeric protein" or "fusion protein" comprises a NgR polypeptide operatively linked to a non-NgR polypeptide. A "NgR polypeptide" refers to a polypeptide having an amino acid sequence corresponding to NgR, whereas a "non-NgR polypeptide" refers to a polypeptide having an amino acid sequence
20 corresponding to a protein that is not homologous to the NgR protein, *e.g.*, a protein that is different from the NgR protein and that is derived from the same or a different organism. Within a NgR fusion protein the NgR polypeptide can correspond to all or a portion of a NgR protein. In one embodiment, a NgR fusion protein comprises at least one biologically active portion of a NgR protein. In another embodiment, a NgR
25 fusion protein comprises at least two biologically active portions of a NgR protein. In yet another embodiment, a NgR fusion protein comprises at least three biologically active portions of a NgR protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the NgR polypeptide and the non-NgR polypeptide are fused in-frame to each other. The non-NgR polypeptide can be fused to the N-
30 terminus or C-terminus of the NgR polypeptide.

For example, in one embodiment a NgR fusion protein comprises a NgR domain operably linked to the extracellular domain of a second protein. Such fusion

- 47 -

proteins can be further utilized in screening assays for compounds which modulate NgR activity (such assays are described in detail below).

For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the
5 desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide.

In another embodiment, the fusion protein is a NgR protein containing a heterologous signal sequence at its N-terminus. For example, the native NgR signal sequence (*i.e.*, amino acids 1-30 of SEQ ID NO:2 and amino acids 1-40 of SEQ ID
10 NO:4) can be removed and replaced with a signal sequence from another protein. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion NgR can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is a NgR-immunoglobulin fusion protein in which the NgR sequences comprising one or more domains are fused to
15 sequences derived from a member of the immunoglobulin protein family. The NgR-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between NgR ligand and a NgR protein on the surface of a cell, to thereby suppress NgR-mediated signal transduction *in vivo*. NgR-immunoglobulin fusion proteins can be
20 used to affect the bioavailability of a NgR cognate ligand. Inhibition of the NgR ligand/NgR interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the NgR-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NgR antibodies in a subject, to
25 purify NgR ligands, and in screening assays to identify molecules that inhibit the interaction of NgR with NgR ligand.

A NgR chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional
30 techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining and

- 48 -

enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can
5 subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel *et al.* (Eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A NgR-encoding nucleic acid can be cloned into such an expression vector such that the fusion
10 moiety is linked in-frame to the NgR protein.

Variants resulting from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of NgR is/are fused to another polypeptide.

15 In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a NgR polypeptide are removed. Deletions can be effected at one or both termini of the NgR polypeptide, or with removal of one or more non-terminal amino acid residues of NgR. Deletion variants, therefore, include all fragments of a NgR polypeptide.

20 The invention also embraces polypeptide fragments of the sequence set forth in SEQ ID NO:2, 4 or 14 wherein the fragments maintain biological (*e.g.*, ligand binding and/or intracellular signaling) immunological properties of a NgR polypeptide. Fragments comprising at least 4, 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of SEQ ID NO:2, 4 or 14 are contemplated by the invention. Preferred
25 polypeptide fragments display antigenic properties unique to, or specific for, human NgR and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of NgR
30 polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a NgR polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the

- 49 -

invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 2, 3, or 4 below.

Table 1.

5	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁	G, R, M	
	X ₂	A, D, C	
	X ₃	V, T	
10	X ₄	N, P, S	
	X ₅	E, A, S	
	X ₆	nothing, K	nothing
	X ₇	V, M, P	
	X ₈	T, V	V
15	X ₉	Q, P	Q
	X ₁₀	Q, A	Q
	X ₁₁	Q, H, N	
	X ₁₂	G, N	N
	X ₁₃	L, F	F
20	X ₁₄	Q, A, S	
	X ₁₅	A, S	
	X ₁₆	V, I	
	X ₁₇	V, T, E, L	
	X ₁₈	S, G	
25	X ₁₉	L, I	
	X ₂₀	A, E, V, P	
	X ₂₁	A, S, D	
	X ₂₂	S, T	
	X ₂₃	Q, E	

- 50 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₂₄	IVL	
	X ₂₅	Q,H	Q
	X ₂₆	N,G	N
	X ₂₇	R,L	
5	X ₂₈	T,G,R,S	
	X ₂₉	F,L,T,H	
	X ₃₀	L,V	L
	X ₃₁	Q,R,P	
	X ₃₂	Q,P,A	P
10	X ₃₃	G,A	G
	X ₃₄	H,T,S	
	X ₃₅	S,G,R	
	X ₃₆	P,S,A	
	X ₃₇	C, nothing	nothing
15	X ₃₈	R, nothing	nothing
	X ₃₉	A,N	
	X ₄₀	M,L	
	X ₄₁	V,L,T	
	X ₄₂	T, I	T
20	X ₄₃	L,I	
	X ₄₄	Y,F,H	
	X ₄₅	N,V	N
	X ₄₆	I,L	
	X ₄₇	T,S,A	
25	X ₄₈	F,Y,T,R	
	X ₄₉	A,H,Y,D	

- 51 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₅₀	P,A	P
	X ₅₁	N,S,G,A	
	X ₅₂	T,A	T
	X ₅₃	E,R,T	
5	X ₅₄	G,H	
	X ₅₅	F,L	
	X ₅₆	V,Q,H	
	X ₅₇	H,A,L	
	X ₅₈	E,Q	E
10	X ₅₉	G,S	G
	X ₆₀	R,A	R
	X ₆₁	Q,H	Q
	X ₆₂	R, H	H
	X ₆₃	T,S	
15	X ₆₄	L,V	L
	X ₆₅	A,E,D	
	X ₆₆	E,D,A	
	X ₆₇	Q,H	Q
	X ₆₈	V,E,G	
20	X ₆₉	K,R	
	X ₇₀	H,Q	
	X ₇₁	A,S,T	
	X ₇₂	Y,H	
	X ₇₃	Y,D	Y
25	X ₇₄	K,R	
	X ₇₅	G,Q	

- 52 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₇₆	S,Q	S
	X ₇₇	A,S,E	
	X ₇₈	P,G	P
	X ₇₉	A,G,P	
5	X ₈₀	G,N	
	X ₈₁	I,V,L	
	X ₈₂	G,R	
	X ₈₃	H,V,A	
	X ₈₄	S,A	S
10	X ₈₅	D,E	
	X ₈₆	H,S,A	
	X ₈₇	I,L	
	X ₈₈	E,L,Q	
	X ₈₉	Y,H,A	
15	X ₉₀	Q,P	Q
	X ₉₁	D, N	
	X ₉₂	I, L, T	
	X ₉₃	V,A,R	
	X ₉₄	V,A,G	
20	X ₉₅	S,T	S
	X ₉₆	K,R	
	X ₉₇	L,I	L
	X ₉₈	W,R,S	
	X ₉₉	S,L	
25	X ₁₀₀	L,V	L
	X ₁₀₁	G,T,P	

- 53 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₀₂	Q,P,E	
	X ₁₀₃	G,H,R	
	X ₁₀₄	I,T,V,A	
	X ₁₀₅	V,G,H	
5	X ₁₀₆	N,S	
	X ₁₀₇	E,G,Q	
	X ₁₀₈	Q,R	
	X ₁₀₉	L,V	
	X ₁₁₀	Q,A	
10	X ₁₁₁	W,G,H	
	X ₁₁₂	H,R,P	
	X ₁₁₃	K,A,H	
	X ₁₁₄	H,R	
	X ₁₁₅	D,G	
15	X ₁₁₆	H,R,S,G	
	X ₁₁₇	T,M	
	X ₁₁₈	T,I	
	X ₁₁₉	F,Y	
	X ₁₂₀	N,A	
20	X ₁₂₁	S,N	
	X ₁₂₂	T,A,S	
	X ₁₂₃	E,S,A	
	X ₁₂₄	Q,P	
	X ₁₂₅	G,T	
25	X ₁₂₆	D,E	D
	X ₁₂₇	C,A	

- 54 -

	X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
	X ₁₂₈	P,D	
	X ₁₂₉	V,G,P,R	
	X ₁₃₀	A,S	
	X ₁₃₁	E,Q	Q
5	X ₁₃₂	F,Y	F
	X ₁₃₃	G,A,D	
	X ₁₃₄	A,P	
	X ₁₃₅	D,A,V	
	X ₁₃₆	G,D	
10	X ₁₃₇	A,E	
	X ₁₃₈	S,P	
	X ₁₃₉	E,A	
	X ₁₄₀	L,F	
	X ₁₄₁	R,Q	
15	X ₁₄₂	R,K	R
	X ₁₄₃	R,K	R
	X ₁₄₄	F,A	
	X ₁₄₅	G,V	
	X ₁₄₆	A,D,E	
20	X ₁₄₇	T,P	
	X ₁₄₈	A,V,S	
	X ₁₄₉	T,S,L	
	X ₁₅₀	E,G,P,Q	
	X ₁₅₁	L,E,R	
25	X ₁₅₂	R,L	R
	X ₁₅₃	G, D	

- 55 -

X_{aa}# (based on a NTLRRCT domain)	Column I (R1, R2, R3)	Column II (R2+R3 only)
X ₁₅₄	Q,H,A	
X ₁₅₅	Q,R	
X ₁₅₆	K,R	
X ₁₅₇	L,A,R	
X ₁₅₈	R,A	R
X ₁₅₉	V,A,E	
X ₁₆₀	E,A,N	
X ₁₆₁	F,L	F
X ₁₆₂	R,Q	
X ₁₆₃	N,A,G	

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 2 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table 2**Conservative Substitutions I**

SIDE CHAIN CHARACTERISTIC	AMINO ACID
Aliphatic	
Non-polar	G A P I L V

- 56 -

Polar - uncharged

C S T M
N Q

Polar - charged

D E
K R

Aromatic

H F W Y

Other

N Q D E

5 Alternatively, conservative amino acids can be grouped as described in Lehninger, [BIOCHEMISTRY, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 3, immediately below.

Table 3**Conservative Substitutions II**

10

**SIDE CHAIN
CHARACTERISTIC****AMINO ACID**

Non-polar (hydrophobic)

A. Aliphatic:

A L I V P

B. Aromatic:

F W

15

C. Sulfur-containing:

M

D. Boderline:

G

Uncharged-polar

A. Hydroxyl:

S T Y

B. Amides:

N Q

C. Sylfhydryl:

C

20

D. Boderline:

G

Positively Charged (Basic):

K R H

Negatively Charged (Acidic):

D E

25 As still another alternative, exemplary conservative substitutions are set out in Table 4, below.

Table 4**Conservative Substitutions III****Original
Residue****Exemplary Substitution**

30

Ala (A)

Val, Leu, Ile

Arg (R)

Lys, Gln, Asn

Asn (N)

Gln, His, Lys, Arg

Asp (D)

Glu

Cys (C)

Ser

Gln (Q)

Asn

35

Glu (E)

Asp

- 57 -

	His (H)	Asn, Gln, Lys, Arg
	Ile (I)	Leu, Val, Met, Ala, Phe,
	Leu (L)	Ile, Val, Met, Ala, Phe
5	Lys (K)	Arg, Gln, Asn
	Met (M)	Leu, Phe, Ile
	Phe (F)	Leu, Val, Ile, Ala
	Pro (P)	Gly
	Ser (S)	Thr
	Thr (T)	Ser
10	Trp (W)	Tyr
	Tyr (Y)	Trp, Phe, Thr, Ser
	Val (V)	Ile, Leu, Met, Phe, Ala

In addition, amino acid residues that are conserved among family members of the NgR proteins of the present invention, as indicated by the alignment presented herein, are also predicted to be particularly unamenable to alteration. For example, NgR proteins of the present invention can contain at least one domain that is a typically conserved region in NgRs. Examples of these conserved domains include, *e.g.*, leucine-rich repeat domain. Amino acid residues that are not conserved or are only semi-conserved among members of the NgR proteins may be readily amenable to alteration.

Full-length NgRs have an LRR region characterized by the amino acid consensus sequence shown in SEQ ID NO: 19. At least some full-length NgRs also include a CT signaling (CTS) domain and a GPI domain.

The NgR domain designations used herein are defined as follows:

Domain	hNgR1 SEQ ID: 5	mNgR1 SEQ ID NO:17	hNgR2 SEQ ID: 2	hNgR3 SEQ ID: 14	mNgR3 SEQ ID: 4
Signal Seq.	1-26	1-26	1-30	—	1-40
LRRNT	27-56	27-56	31-59	—	41-69
LRR1	57-81	57-81	60-82	5-27	70-92
LRR2	82-105	82-105	83-106	28-51	93-106
LRR3	106-130	106-130	107-131	52-76	106-141
LRR4	131-154	131-154	132-155	77-100	142-165

- 58 -

5	LRR5	155-178	155-178	156-179	101-124	166-189
	LRR6	179-202	179-202	180-203	125-148	190-213
	LRR7	203-226	203-226	204-227	149-172	214-237
	LRR8	227-250	227-250	228-251	173-196	238-261
	LRRCT	260-309	260-309	261-310	206-255	271-320
	CTS (CT Signaling)	310-445	310-445	311-395	256-396	321-438
	GPI	446-473	456-473	396-420	370-392	439-462

10

In some embodiments of the invention, the above domains are modified. Modification can be in a manner that preserves domain functionality. Modification can include addition, deletion or substitution of certain amino acids. Exemplary modifications include conservative amino acid substitutions. Preferably such substitutions number 20 or fewer per 100 residues. More preferably, such substitutions number 10 or fewer per 100 residues. Further exemplary modifications include addition of flanking sequences of up to five amino acids at the N terminus and/or C terminus of one or more of the domains.

20 In some embodiments, the isolated nucleic acid molecule encodes a polypeptide at least about 70%, 80%, 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2, 4 or 14.

Mutations can be introduced into SEQ ID NOS:1, 3 or 13 by standard techniques, e.g., site-directed mutagenesis and PCR-mediated mutagenesis. Conservative amino acid substitutions can be made at one or more amino acid residues predicted to be non-essential. Alternatively, mutations can be introduced randomly along a NgR coding sequence. This can be accomplished, e.g., by saturation mutagenesis. The resulting mutants can be screened for NgR biological activity. Biological activities of NgR may include but are not limited to: (1) protein:protein interactions, e.g., with other NgRs or other cell-surface proteins involved in Nogo-related signaling; (2) complex formation with a NgR ligand; (3) binding to an anti-NgR antibody.

- 59 -

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues or organs. Similarly, the invention further embraces NgR polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol or polypropylene glycol. Variants that display ligand binding properties of native NgR and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

Chemically modified NgR polypeptide compositions in which the NgR polypeptide is linked to a polymer are included within the scope of the present invention. The polymer may be water soluble to prevent precipitation of the protein in an aqueous environment, such as a physiological environment. Suitable water-soluble polymers may be selected from the group consisting of, for example, polyethylene glycol (PEG), monomethoxypolyethylene glycol, dextran, cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, polypropylene glycol homopolymers, a polypropylene oxide/ethylene oxide copolymer polyoxyethylated polyols (*e.g.* glycerol) and polyvinyl alcohol. The selected polymer is usually modified to have a single reactive group, such as an active ester for acylation or an aldehyde for alkylation, so that the degree of polymerization may be controlled. Polymers may be of any molecular weight, and may be branched or unbranched, and mixtures of such polymers may also be used. When the chemically modified NgR polymer is destined for therapeutic use, pharmaceutically acceptable polymers will be selected for use.

When the polymer is to be modified by an acylation reaction, the polymer should have a single reactive ester group. Alternatively, if the polymer is to be modified by reductive alkylation, the polymer should have a single reactive aldehyde

- 60 -

group. A preferred reactive aldehyde is polyethylene glycol propionaldehyde, which is water stable, or mono Cl-CIO alkoxy or aryloxy derivatives thereof (see U.S. Patent No. 5,252,714, incorporated by reference herein in its entirety).

Pegylation of NgR polypeptides may be carried out by any of the pegylation reactions known in the art, as described, for example, in the following references: *Focus on Growth Factors* 3, 4-10 (1992); EP 0 154 316; and EP 0 401 384 (each of which is incorporated by reference herein in its entirety). Preferably, the pegylation is carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer). A preferred water-soluble polymer for pegylation of polypeptides such as NgR is polyethylene glycol (PEG). As used herein, "polyethylene glycol" is meant to encompass any of the forms of PEG that have been used to derivatize other proteins, such as mono (Cl-CIO) alkoxy- or aryloxy-polyethylene glycol.

Chemical derivatization of NgR polypeptides may be performed under any suitable conditions used to react a biologically active substance with an activated polymer molecule. Methods for preparing pegylated NgR polypeptides will generally comprise the steps of (a) reacting the polypeptide with polyethylene glycol, such as a reactive ester or aldehyde derivative of PEG, under conditions whereby NgR polypeptide becomes attached to one or more PEG groups, and (b) obtaining the reaction products. It will be apparent to one of ordinary skill in the art to select the optimal reaction conditions or the acylation reactions based on known parameters and the desired result.

Pegylated and other polymer:NgR polypeptides may generally be used to treat conditions that may be alleviated or modulated by administration of the NgR polypeptides described herein. However, the chemically-derivatized polymer:NgR polypeptide molecules disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the nonderivatized molecules. The NgR polypeptides, fragments thereof, variants and derivatives, may be employed alone, together, or in combination with other pharmaceutical compositions. The cytokines, growth factors, antibiotics, antiinflammatories and/or chemotherapeutic agents as is appropriate for the indication being treated.

- 61 -

The present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient or medium.

5 Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil and cocoa butter.

10 Variants that display ligand binding properties of native NgR and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant NgR activity.

15 With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences which encode NgR from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is
20 incorporated herein by reference in its entirety.

For example, DNA that encodes NgR may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the NgR gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in
25 accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.* (1989) above.

A nucleic acid molecule comprising any of the NgR nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the
30 nucleotide sequences provided herein. See U.S. Patent Nos. 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that

- 62 -

sequence has not been previously purified and is present only in a single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase-mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, METHODS IN ENZYMOLOGY 152 Academic Press, San Diego, CA, which is incorporated herein by reference in its entirety.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

15 **Antibodies**

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for NgR or fragments thereof. Preferred antibodies of the invention are human antibodies which are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')₂, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind NgR polypeptides exclusively (*i.e.*, are able to distinguish NgR polypeptides from other known NgR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between NgR and such polypeptides).

The antigenic peptide of NgR comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2, 4 or 14 and encompasses an epitope of

- 63 -

NgR such that an antibody raised against the peptide forms a specific immune complex with NgR. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. Preferred epitopes
5 encompassed by the antigenic peptide are regions of NgR that are located on the surface of the protein, *e.g.*, hydrophilic regions.

It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and,
10 in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* in ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments
15 of the NgR polypeptides of the invention are also contemplated, provided that the antibodies are specific for NgR polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by injection with the native
20 protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed NgR protein or a chemically synthesized NgR polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*,
25 aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum* or similar immunostimulatory agents. If desired, the antibody molecules directed against NgR can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such
30 as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition," as used herein, refers to a population of antibody molecules that contain only one species

- 64 -

of an antigen binding site capable of immunoreacting with a particular epitope of NgR. A monoclonal antibody composition thus typically displays a single binding affinity for a particular NgR protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular NgR protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler and Milstein (1975) *Nature* 256, 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor *et al.*, (1983) *Immunol. Today* 4, 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole *et al.*, (1985) in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote *et al.*, (1983) *Proc. Natl. Acad. Sci. USA* 80, 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole *et al.*, (1985), above).

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a NgR protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of Fab expression libraries (see *e.g.*, Huse *et al.*, (1989) *Science* 246, 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a NgR protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity. Antibody fragments that contain the idiotypes to a NgR protein may be produced by techniques known in the art including, but not limited to: (i) an F(ab')₂ fragment produced by pepsin digestion of an antibody molecule; (ii) an Fab fragment generated by reducing the disulfide bridges of an F(ab')₂ fragment; (iii) an Fab fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

- 65 -

Additionally, recombinant anti-NgR antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be

5 produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.*,

10 (1988) *Science* 240, 1041-1043; Liu *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 3439-3443; Liu *et al.*, (1987) *J. Immunol.* 139, 3521-3526; Sun *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 84, 214-218; Nishimura *et al.*, (1987) *Cancer Res.* 47, 999-1005; Wood *et al.*, (1985) *Nature* 314, 446-449; Shaw *et al.*, (1988) *J. Natl. Cancer Inst.* 80, 1553-1559; Morrison (1985) *Science* 229, 1202-1207; Oi *et al.*, (1986)

15 *BioTechniques* 4, 214; U.S. Patent. No. 5,225,539; Jones *et al.*, (1986) *Nature* 321, 552-525; Verhoeyan *et al.*, (1988) *Science* 239, 1534; and Beidler *et al.*, (1988) *J. Immunol.* 141, 4053-4060.

In a preferred embodiment of the invention a portion of a NgR is joined to an Fc portion of an antibody to form a NgR/Fc fusion protein. Preferably, the Ig fusion

20 protein is soluble. The NgR/Fc fusion protein may be formed by recombinant techniques as described above. In one embodiment, a portion of a NgR including the entire amino acid sequence of NgR except the C-terminal hydrophobic region is fused to an Fc portion of an antibody. In preferred embodiments, the NgR is a human NgR and the Fc is also human. More preferably, the human Fc portion is derived from an

25 IgG antibody. In other embodiments, the N-terminal signal sequence is omitted. Such antibodies are useful in binding Nogo to prevent Nogo signaling through the NgR.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a

30 specific embodiment, selection of antibodies that are specific to a particular domain of a NgR protein is facilitated by generation of hybridomas that bind to the fragment of a NgR protein possessing such a domain. Antibodies that are specific for one or more

- 66 -

domains within a NgR protein, *e.g.*, domains spanning the above-identified conserved regions of NgRs, or derivatives, fragments analogs or homologs thereof, are also provided herein.

Anti-NgR antibodies may be used in methods known within the art relating to the localization and/or quantitation of a NgR protein (*e.g.*, for use in measuring levels of the NgR protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for NgR proteins, or derivatives, fragments analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-NgR antibody (*e.g.*, monoclonal antibody) can be used to isolate NgR by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-NgR antibody can facilitate the purification of natural NgR from cells and of recombinantly produced NgR expressed in host cells. Moreover, an anti-NgR antibody can be used to detect NgR protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the NgR protein. Anti-NgR antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering

- 67 -

to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Another aspect of the invention is directed to anti-idiotypic antibodies and
5 anti-anti-idiotypic antibodies. An anti-idiotypic antibody is an antibody that recognizes determinants of another antibody (a target antibody). Generally, the anti-idiotypic antibody recognizes determinants of the antigen-binding site of the target antibody. Typically, the target antibody is a monoclonal antibody. An anti-idiotypic antibody is generally prepared by immunizing an animal (particularly, mice) of the same species
10 and genetic type as the source of the target monoclonal antibody, with the target monoclonal antibody. The immunized animal mounts an immune response to the idiotypic determinants of the target monoclonal antibody and produces antibodies against the idiotypic determinants of the target monoclonal antibody. Antibody-producing cells, such as splenic cells, of the immunized animal may be used
15 to generate anti-idiotypic monoclonal antibodies. Furthermore, an anti-idiotypic antibody may also be used to immunize animals to produce anti-anti-idiotypic antibodies. These immunized animals may be used to generate anti-anti-idiotypic monoclonal antibodies using standard techniques. The anti-anti-idiotypic antibodies may bind to the same epitope as the original, target monoclonal antibody used to
20 prepare the anti-idiotypic antibody. The anti-anti-idiotypic antibodies represent other monoclonal antibodies with the same antigen specificity as the original target monoclonal antibody.

If the binding of the anti-idiotypic antibody with the target antibody is inhibited by the relevant antigen of the target antibody, and if the anti-idiotypic antibody induces
25 an antibody response with the same specificity as the target antibody, it mimics the antigen of the target antibody. Such an anti-idiotypic antibody is an "internal image anti-idiotype" and is capable of inducing an antibody response as if it were the original antigen. (Bona and Kohler (1984) ANTI-IDIOTYPIC ANTIBODIES AND INTERNAL IMAGE, IN MONOCLONAL AND ANTI-IDIOTYPIC ANTIBODIES: PROBES FOR RECEPTOR STRUCTURE
30 AND FUNCTION, Venter J.C. *et al.* (Eds), Alan R. Liss, New York, NY, pp 141-149, 1984). Vaccines incorporating internal image anti-idiotype antibodies have been shown to induce protective responses against viruses, bacteria, and parasites (Kennedy

- 68 -

et al., (1986) 232, 220-223; 1047; McNamara *et al.*, (1985) *Science* 226, 1325-1326). Internal image anti-idiotypic antibodies have also been shown to induce immunity to tumor related antigens (Raychauhuri *et al.*, (1986) *J. Immunol.* 137, 1743-1749; Raychauhuri *et al.*, (1987) *J. Immunol.* 139, 3902-3910; Bhattacharya-Chatterjee *et al.*, (1987) *J. Immunol.* 139, 1354-1360; Bhattacharya-Chatterjee *et al.*, (1988) *J. Immunol.* 141, 1398-1403; Herlyn. *et al.* (1989) *Intern. Rev. Immunol.* 4, 347-357; Chen *et al.* (1990) *Cell Imm. Immunother. Cancer* 351-359; Herlyn *et al.*, (1991) *in vivo* 5, 615-624; Furuya *et al.* (1992) *AntiCancer Res.* 12, 27-32; Mittelman, A. *et al.* (1992) *Proc. Natl. Acad. Sci., USA* 89, 466-470; Durrant. *et al.*, (1994) *Cancer Res.* 54, 4837-4840; Mittelman. *et al.* (1994) *Cancer Res.* 54, 415-421; Schmitt. *et al.* (1994) *Hybridoma* 13, 389-396; Chakrobarty. *et al.* (1995) *J. Immunother.* 18, 95-103; Chakrobarty. *et al.* (1995) *Cancer Res.* 55, 1525-1530; Foon, K. A. *et al.* (1995) *Clin. Cancer Res.* 1, 1205-1294; Herlyn *et al.* (1995) *Hybridoma* 14, 159-166; Sclebusch *et al.* (1995) *Hybridoma* 14, 167-174; Herlyn. *et al.* (1996) *Cancer Immunol Immunother.* 43, 65-76).

Anti-idiotypic antibodies for NgR may be prepared, for example, by immunizing an animal, such as a mouse, with a immunogenic amount of a composition comprising NgR2 (SEQ ID NO:2), NgR3 (SEQ ID NOs:4 or 14), or immunogenic portion thereof, containing at least one antigenic epitope of NgR. The composition may also contain a suitable adjuvant, and any carrier necessary to provide immunogenicity. Monoclonal antibodies recognizing NgR may be prepared from the cells of the immunized animal as described above. A monoclonal antibody recognizing an epitope of NgR is then selected and used to prepare a composition comprising an immunogenic amount of the anti-NgR monoclonal antibody. Typically, a 25 to 200 μ g dose of purified anti-NgR monoclonal would be sufficient in a suitable adjuvant.

Animals may be immunized 2-6 times at 14 to 30 day intervals between doses. Typically, animals are immunized by any suitable route of administration, such as intraperitoneal, subcutaneous, intravenous or a combination of these. Anti-idiotypic antibody production may be monitored during the immunization period using standard immunoassay methods. Animals with suitable titers of antibodies reactive with the target monoclonal antibodies may be reimmunized with the monoclonal antibody used as the immunogen three days before harvesting the antibody producing cells.

- 69 -

Preferably, spleen cells are used, although other antibody producing cells may be selected. Antibody-producing cells are harvested and fused with myeloma cells to produce *Hybridomas*, as described above, and suitable anti-idiotypic antibody-producing cells are selected.

5 Anti-anti-idiotypic antibodies are produced by another round of immunization and *Hybridoma* production by using the anti-idiotypic monoclonal antibody as the immunogen.

 Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of NgR), diagnostic purposes to detect or quantitate NgR, and
10 purification of NgR. Therefore, kits comprising an antibody of the invention for any of the purposes described herein are also comprehended.

Kits

 The present invention is also directed to kits, including pharmaceutical kits.
15 The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well appropriate controls, such as positive and/or negative controls. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.
20 For example, the kit can comprise: a labeled compound or agent capable of detecting NgR protein or mRNA in a biological sample; means for determining the amount of NgR in the sample; and means for comparing the amount of NgR in the sample with a standard. The compound or agent can be packaged in a suitable container.

25 Screening Assays

 The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a NgR polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein NgR polypeptides are
30 immobilized and cell-based assays. Identification of binding partner compounds of NgR polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with NgR normal and aberrant biological activity.

- 70 -

The invention also provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules (*e.g.*, molecules of less than 1,000 Daltons) or other drugs) that bind to NgR proteins or have a stimulatory or inhibitory effect on, for example, NgR expression or NgR activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a NgR protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12, 145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90, 6909; Erb *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91, 11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37, 2678; Cho *et al.*, (1993) *Science* 261, 1303; Carrell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2059; Carell *et al.*, (1994) *Angew Chem. Int. Ed. Engl.* 33, 2061; and Gallop *et al.*, (1994) *J. Med. Chem.* 37, 1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *BioTechniques* 13, 412-421), or on beads (Lam (1991) *Nature* 354, 82-84), on chips (Fodor (1993) *Nature* 364, 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, above), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89, 1865-1869) or on phage (Scott and Smith (1990) *Science* 249, 386-390; Devlin (1990) *Science* 249, 404-406; Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 6378-6382; Felici (1991) *J. Mol. Biol.* 222, 301-310; Ladner, above).

1. Cell-based Assays

The invention also provides cell-based assays to identify binding partner compounds of a NgR polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a NgR polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the NgR polypeptide. In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the NgR protein or biologically active portion thereof.

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NgR protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a NgR protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NgR protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NgR protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NgR protein or a biologically active portion thereof, on the cell surface with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the test compound to preferentially bind to NgR or a biologically active portion thereof as compared to the known compound.

- 72 -

Determining the ability of the test compound to modulate the activity of NgR or a biologically active portion thereof can be accomplished, for example, by determining the ability of the NgR protein to bind to or interact with a NgR target molecule. As used herein, a "target molecule" is a molecule with which a NgR protein
5 binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a NgR protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A NgR target molecule can be a non-NgR molecule or a NgR protein or polypeptide of the present invention. In one embodiment, a NgR
10 target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound NgR molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling
15 molecules with NgR. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant NgR products, NgR
20 variants, or preferably, cells expressing such products. Binding partners are useful for purifying NgR products and detection or quantification of NgR products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of NgR, especially those activities involved in signal transduction.

25

2. Cell-free Assays

(a) Direct binding:

The invention includes several assay systems for identifying NgR binding partners. In solution assays, methods of the invention comprise the steps of (a)
30 contacting a NgR polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the NgR polypeptide. Identification of the compounds that bind the NgR polypeptide can be achieved by isolating the NgR

- 73 -

polypeptide/binding partner complex and separating the binding partner compound from the NgR polypeptide. An additional step of characterizing the physical, biological and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the NgR polypeptide/binding partner complex is isolated using an antibody immunospecific for either the NgR polypeptide or the candidate binding partner compound.

In still other embodiments, either the NgR polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the NgR polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG[®] tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

15

(b) Immobilized NgR

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized NgR polypeptide, or a biologically active fragment thereof with a candidate binding partner compound and (b) detecting binding of the candidate compound to the NgR polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of NgR is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Binding of a test compound to NgR, or interaction of NgR with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, and not by way of limitation, GST-NgR fusion proteins or GST-target fusion proteins can be

30

- 74 -

adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or NgR protein, and the mixture is incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, and the complexes determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of NgR binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either NgR or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated NgR or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NgR or target molecules, but which do not interfere with binding of the NgR protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NgR trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NgR or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NgR or target molecule.

Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, (v) determining the activity of the NgR, as well as other techniques well known and routinely practiced in the art.

- 75 -

Determining the activity of the target molecule, for example, may be accomplished by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising
5 a NgR-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

(c) Competition experiments

10 In yet another embodiment, the assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining
15 the ability of the test compound to preferentially bind to NgR or biologically active portion thereof as compared to the known compound.

In yet another embodiment, the cell-free assay comprises contacting the NgR protein or biologically active portion thereof with a known compound which binds NgR to form an assay mixture, contacting the assay mixture with a test compound, and
20 determining the ability of the test compound to interact with a NgR protein, wherein determining the ability of the test compound to interact with a NgR protein comprises determining the ability of the NgR protein to modulate the activity of a NgR target molecule.

The cell-free assays of the present invention are amenable to use of both the
25 soluble form or the membrane-bound form of NgR. In the case of cell-free assays comprising the membrane-bound form of NgR, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of NgR is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide,
30 decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_m, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-

- 76 -

propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

Modulators

5 Agents that modulate (*i.e.*, increase, decrease, or block) NgR activity or expression may be identified by incubating a putative modulator with a cell containing a NgR polypeptide or polynucleotide and determining the effect of the putative modulator on NgR activity or expression. The selectivity of a compound that modulates the activity of NgR can be evaluated by comparing its effects on NgR to its
10 effect on other NgR compounds. Selective modulators may include, for example, antibodies and other proteins, peptides or organic molecules which specifically bind to a NgR polypeptide or a NgR-encoding nucleic acid. Modulators of NgR activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant NgR activity is involved. NgR polynucleotides, polypeptides and
15 modulators may be used in the treatment of such diseases and conditions associated with demyelination. NgR polynucleotides and polypeptides, as well as NgR modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations
20 including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the NgR polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A
25 modulator that increases binding between the NgR polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the NgR polypeptide and the binding partner compound is described as an inhibitor.

In another embodiment, modulators of NgR expression may be identified in a
30 method wherein a cell is contacted with a candidate compound and the expression of NgR mRNA or protein in the cell is determined. The level of expression of NgR mRNA or protein in the presence of the candidate compound is compared to the level

- 77 -

of expression of NgR mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NgR expression based on this comparison. For example, when expression of NgR mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NgR mRNA or protein expression. Alternatively, when expression of NgR mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of NgR mRNA or protein expression. The level of NgR mRNA or protein expression in the cells can be determined by methods described herein for detecting NgR mRNA or protein.

High Throughput Screening

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, etc.) of a NgR polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate NgR receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the NgR polypeptide.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either NgR or nucleic acid molecules encoding NgR, comprising contacting NgR, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds NgR or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, Ausubel *et al.* (Eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in

its entirety. The NgR proteins, for example, can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, (1993) *Cell* 72, 223-232; Madura *et al.*, (1993) *J. Biol. Chem.* 268, 12046-12054; Bartel *et al.*, (1993) *BioTechniques* 14, 920-924; Iwabuchi *et al.*, (1993) *Oncogene* 8, 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact
5 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with NgR ("NgR-binding proteins" or "NgR-bp") and modulate NgR activity. Such NgR-binding proteins are also likely to be involved in the propagation of signals by the NgR proteins as, for example, upstream or downstream elements of the NgR pathway.

Other assays may be used to identify specific ligands of a NgR receptor,
10 including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in
15 Fields *et al.*, (1989) *Nature* 340, 245-246, and Fields *et al.*, (1994) *Trends Genet.* 10, 286-292, both of which are incorporated herein by reference. The two-hybrid system is a genetic assay based on the modular nature of most transcription factors used for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues
20 critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a
25 reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate
30 transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the

- 79 -

noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a NgR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal. The compounds to be screened include (which may include compounds that are suspected to bind NgR, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biological or chemical origin.

The function of the NgR gene product is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay is useful to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a NgR receptor, or fragment thereof, a fusion polynucleotide encoding both a NgR receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of

- 80 -

the ligand to the target protein can be determined by any method which distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.* (1997) *Anal. Chem.* 69:1683-1691, incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I , ^{35}S , ^{32}P , ^{33}P , ^3H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The NgR polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between NgR and its substrate caused by the compound being tested.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of NgR comprising contacting NgR with a compound, and determining whether the compound modifies activity of NgR. The activity in the presence of the test compared is measured to the

- 81 -

activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited activity.

The present invention is particularly useful for screening compounds by using NgR in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate NgR activity), but are not limited to, extracellular, intracellular, biologic or chemical origin.

10 The NgR polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between NgR and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Nogo-R and its substrate caused by the

15 compound being tested.

The activity of NgR polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands. Alternatively, the activity of the NgR can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides,

20 neurotransmitters, nucleotides, lipids, odorants and photons. Alternatively, the activity of the NgR can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of NgR activity may alter a NgR receptor function, such as a binding property of a receptor or an activity. In various embodiments of the method, the assay

25 may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [³⁵S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca²⁺ concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (e.g., using [³H]-arachidonic acid) and an assay for

30 extracellular acidification rates, as well as other binding or function-based assays of NgR activity that are generally known in the art. NgR activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those

- 82 -

skilled in the art. Biological activities of NgR receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of NgRs known in the art. Non-limiting examples of NgR activities include transmembrane signaling of various forms, which
5 may involve phosphatidylinositol (PI) association and/or the exertion of an influence over PI; another exemplary activity of NgRs is the binding of accessory proteins or polypeptides that differ from known GPI proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural NgR receptor ligands, peptide and non-peptide allosteric effectors of NgR receptors, and peptides that may
10 function as activators or inhibitors (competitive, uncompetitive and non-competitive) (*e.g.*, antibody products) of NgR receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule
15 libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, *ENZYME ASSAYS: A PRACTICAL APPROACH*, Eisenthal and
20 Danson (Eds.), 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use
25 of radiolabelled ligands in HTS binding assays for drug discovery (see Williams (1991) *Med. Res. Rev.*, 11, 147-184; Sweetnam *et al.*, (1993) *J. Nat. Prod.* 56, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see
30 Hodgson (1992) *Bio/Technology* 10, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg *et al.* (1992) *Trends Pharmacol. Sci.* 13, 95-98), yeast (Pausch (1997) *Trends Biotechnol.* 15, 487-494), several kinds of insect cells (Vanden
5 Broeck (1996) *Int. Rev. Cytol.* 164, 189-268), amphibian cells (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt *et al.* (1997) *Eur. J. Pharmacol.* 334, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

10 In preferred embodiments of the invention, methods of screening for compounds which modulate NgR activity comprise contacting test compounds with NgR and assaying for the presence of a complex between the compound and NgR. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed
15 label is a measure of the ability of the particular compound to bind to NgR.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to NgR is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with NgR and washed. Bound NgR is then
20 detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed NgR can be used for HTS binding assays in
25 conjunction with its defined ligand. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I , ^3H , ^{35}S or ^{32}P , by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.* (1994) *Drug Dev. Res.* 33, 373-398; Rogers (1997) *Drug Discov. Today* 2, 156-160). Radioactive
30 ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate

- 84 -

- bound ligand from unbound ligand (Williams (1991) *Med. Res. Rev.* 11, 147-184; Sweetnam *et al.* (1993) *J. Nat. Prod.* 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama (1998) *Curr. Opin. Drug Disc. Dev.* 1, 85-91 Bossé *et al.* (1998) *J. Biomol. Screening* 3, 285-292). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers (1997) *Drug Discov. Today* 2, 156-160; Hill (1998) *Curr. Opin. Drug Disc. Dev.* 1, 92-97).
- Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch (1997) *Trends in Biotechnol.* 15, 487-494); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy *et al.* (1998) *Cur. Opin. Drug Disc. Dev.* 1, 192-199). Fluorescence changes can also be used to monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder *et al.* (1996) *J. Biomol. Screening* 1, 75-80). Melanophores prepared from *Xenopus laevis* show a ligand-dependent change in pigment organization in response to heterologous NgR activation; this response is adaptable to HTS formats (Jayawickreme *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.
- Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to transduce the signal for the binding of Nogo in membranes prepared from these cells through the putative GPI anchor. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca^{2+} concentration or membrane potential in permanently transfected CHO cells containing

- 85 -

each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (e.g., Pausch (1997), above).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to NgR receptors. In one example, the NgR receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the NgR receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the NgR receptor and its binding partner. Another contemplated assay involves a variation of the di-hybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms that are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Cane *et al.*, *Science* (1998) 282, 63-68. Combinatorial libraries are composed of

- 86 -

large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide,
5 protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers (1997) *Curr. Opin. Biotechnol.* 8, 701-707. Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate
10 activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other
15 than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified NgR gene.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the
20 invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with NgR. Radiolabeled competitive binding studies are described in Lin *et al.*, (1997) *Antimicrob. Agents Chemother.* 41, 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

25 In other embodiments of the invention, the polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are
30 incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of

- 87 -

polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors.

- 5 In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing.
- 10 This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

- Alternatively, compounds may be identified which exhibit similar properties to the ligand for the NgR of the invention, but which are smaller and exhibit a longer half
- 15 time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid
- 20 randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

- Comparison of the protein sequence of the present invention with the
- 25 sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of NgR can be
- 30 designed.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Compositions and Pharmaceutical Compositions

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral or parenteral administration. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's PHARMACEUTICAL SCIENCES, 16th ed., (1980) Osol, A (Ed.), which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include oral and parenteral (*e.g.*, intravenous, intradermal, subcutaneous, inhalation, transdermal (topical), transmucosal and rectal administration). Solutions or suspensions used for parenteral, intradermal or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic

acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a NgR protein or anti-NgR antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of

- 90 -

preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They
5 can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.
10 Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel or corn starch; a lubricant
15 such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable
20 propellant, *e.g.*, a gas such as carbon dioxide or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile
25 salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with
30 conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

- 91 -

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811. It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of routes, *e.g.*, as described in U.S. Patent No. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Patent No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91, 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack or dispenser together with instructions for administration.

- 92 -

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be administered. Therapy is typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

Another aspect of the present invention is the use of the NgR nucleotide sequences disclosed herein for identifying homologs of the Nogo-R, in other animals, including but not limited to humans and other mammals and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more preferably at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with NgR sequences can be identified.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson *Science* (1992) 256, 808-813, which is incorporated herein by reference in its entirety.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing a NgR natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize NgR-associated functions.

- 93 -

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996, and International patent publication number WO 96/22976, published August 1, 1996, both of which are
5 incorporated herein by reference in their entirety, including any drawings, figures or tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition
10 of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

15 Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection
20 methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell
25 composition. For example, toxicity studies can be carried out in a suitable animal model as follows: (1) the compound is administered to mice (an untreated control mouse should also be used); (2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and (3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes
30 versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

- 94 -

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, (1993) *J. Am. Vet. Med. Assoc.* 202:229-249). Representative
5 animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the
10 less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness.
15 Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

NgR mRNA transcripts have been found in the brain and heart. SEQ ID NOs: 1 and/or, 3 will, as detailed above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize NgR and
20 for compounds with potential utility in treating disorders including CNS disorders (*e.g.*, stroke) and degenerative disorders such as those associated with demyelination.

For example, NgR receptor activation may mediate the prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, *e.g.*, Donovan *et al.*, (1997) *J. Neurosci.* 17, 5316-5326; Turgeon *et al.*, (1998) *J. Neurosci.* 18, 6882-6891; Smith-Swintosky *et al.*, (1997) *J. Neurochem.* 69,
25 1890-1896; Gill *et al.*, (1998) *Brain Res.* 797, 321-327; Suidan *et al.*, (1996) *Semin. Thromb. Hemost.* 22, 125-133.

Pharmacogenomics

Agents, or modulators that have a stimulatory or inhibitory effect on NgR activity (*e.g.*, NgR gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, a disease condition such as a demyelination disorder) associated with aberrant NgR activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NgR protein, expression of NgR nucleic acid or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum (1996) *Clin. Exp. Pharmacol. Physiol.* 23, 983-985 and Linder (1997) *Clin. Chem.* 43, 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2

- 96 -

(NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of NgR protein, expression of NgR nucleic acid, or mutation content of NgR genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a NgR modulator, such as a modulator identified by one of the exemplary screening assays described herein.

25

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NgR (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase NgR gene expression, protein levels or upregulate NgR activity, can be monitored in clinical trials of subjects exhibiting decreased NgR gene

30

- 97 -

expression, protein levels, or downregulated NgR activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NgR gene expression, protein levels, or downregulate NgR activity, can be monitored in clinical trials of subjects exhibiting increased NgR gene expression, protein levels, or
5 upregulated NgR activity. In such clinical trials, the expression or activity of NgR and, preferably, other genes that have been implicated in, for example, a disease or disorder, can be used as a "read out" or markers of the immune responsiveness of a particular cell.

For example, genes, including NgR, that are modulated in cells by treatment
10 with an agent (*e.g.*, compound, drug or small molecule) that modulates NgR activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on demyelination disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of NgR and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a
15 gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced by one of the methods as described herein or by measuring the levels of activity of NgR or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state
20 may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate
25 identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a NgR protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the NgR
30 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the NgR protein, mRNA or genomic DNA in the pre-administration sample with the NgR protein, mRNA or genomic DNA in the post

- 98 -

administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NgR to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration
5 of the agent may be desirable to decrease expression or activity of NgR to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods
10 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
15 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to; (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous
20 insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous recombination (see, *e.g.*, Capecchi (1989) *Science* 244, 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction
25 between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner.
30 Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

- 99 -

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art
5 include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a
10 disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic
15 agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

20 Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a
25 protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity.
30 Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the

- 100 -

agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described
5 herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

10 Gene Therapy

Mutations in the NgR gene that result in loss of normal function of the NgR gene product underlie NgR human disease states. The invention comprehends gene therapy to restore NgR activity to treat those disease states. Delivery of a functional NgR gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors,
15 and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson (1998) *Nature*, supplement to 392(6679):25-20. For additional reviews of gene therapy technology see Friedmann (1989) *Science* 244, 1275-1281; Verma (1990) *Sci. Am.* 68-84; and Miller (1992)
20 *Nature* 357, 455-460. Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the activity of, NgR will be useful in treating disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of NgR.

The present invention provides for both prophylactic and therapeutic methods
25 of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NgR expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that
30 antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a NgR polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a NgR

- 101 -

peptide; (iii) nucleic acids encoding a NgR peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a NgR peptide) are utilized to "knockout" endogenous function of a NgR peptide by homologous recombination (see, *e.g.*, Capecchi (1989), above); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a NgR peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, a NgR peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a NgR peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant NgR expression or activity, by administering to the subject an agent that modulates NgR expression or at least one NgR activity. Subjects at risk for a disease that is caused or contributed to by aberrant NgR expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the NgR aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of NgR aberrancy, for example, a NgR agonist or

- 102 -

NgR antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating NgR expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NgR protein activity associated with the cell. An agent that modulates NgR protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a NgR protein, a peptide, a NgR peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NgR protein activity. Examples of such stimulatory agents include active NgR protein and a nucleic acid molecule encoding NgR that has been introduced into the cell. In another embodiment, the agent inhibits one or more NgR protein activity. Examples of such inhibitory agents include antisense NgR nucleic acid molecules and anti-NgR antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a NgR protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) NgR expression or activity. In another embodiment, the method involves administering a NgR protein or nucleic acid molecule as therapy to compensate for reduced or aberrant NgR expression or activity.

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figure. Such modifications are intended to fall within the scope of the appended claims.

The following Table 5 contains the sequences of exemplary polynucleotides and polypeptides of the invention.

TABLE 5

The following DNA sequence NgR2 <SEQ ID NO. 1> was identified in humans:

```

5  ATGCTGCCCCGGGCTCAGGCGCCTGCTGCAAGCTCCCGCCTCGGCCTGCCTCCTGCTGATG
   CTCCTGGCCCTGCCCCTGGCGGCCCCAGCTGCCCCATGCTCTGCACCTGCTACTCATCC
   CCGCCCACCGTGAGCTGCCAGGCCAACAACTTCTCCTCTGTGCCGCTGTCCCTGCCACCC
   AGCACTCAGCGACTCTTCCTGCAGAACAACTCATCCGCACGCTGCGGCCAGGCACCTTT
   GGGTCCAACCTGCTCACCTGTGGCTCTTCTCCAACAACCTCTCCACCATCTACCCGGGC
10  ACTTTCGGCCACTTGCAAGCCCTGGAGGAGCTGGACCTCGGTGACAACCGGCACCTGCGC
   TCGCTGGAGCCCGACACCTTCCAGGGCCTGGAGCGGCTGCAGTCGCTGCATTGTACCGC
   TGCCAGCTCAGCAGCCTGCCCGGCAACATCTTCCGAGGCCTGGTCAGCCTGCAGTACCTC
   TACCTCCAGGAGAACAGCCTGCTCCACCTACAGGATGACTTGTTTCGGGACCTGGCCAAC
   CTGAGCCACCTCTTCTCCACGGGAACCGCCTGCGGCTGCTCACAGAGCACGTGTTTCGC
15  GGCCTGGGCAGCCTGGACCGGCTGCTGCTGCACGGGAACCGGCTGCAGGGCGTGACCCGC
   GCGGCCTTCCGCGGCCTCAGCCGCCTCACCATCCTCTACCTGTTCAACAACAGCCTGGCC
   TCGCTGCCCCGGCGAGGCGCTCGCCGACCTGCCCTCGCTCGAGTTCCTGCGGCTCAACGCT
   AACCCTGGGCGTGCGACTGCCGCGCGCGGCCGCTCTGGGCCTGGTTCAGCGCGCGCGC
   GTGTCCAGCTCCGACGTGACCTGCGCCACCCCCCGAGCGCCAGGGCCGAGACCTGCGC
20  GCGCTCCGCGAGGCCGACTTCCAGGCGTGTCCGCCCGCGGCACCCACGCGGCCGGGCAGC
   CGCGCCCGCGCAACAGCTCCTCCAACCACTGTACGGGGTGGCCGAGGCCGGGGCGCCC
   CCAGCCGATCCCTCCACCCTCTACCGAGATCTGCCTGCCGAAGACTCGCGGGGGCGCCAG
   GGCGGGGACGCGCCTACTGAGGACGACTACTGGGGGGGTACGGGGGTGAGGACCAGCGA
   GGGGAGCAGATGTGCCCCGGCGCTGCCTGCCAGGCGCCCCCGGACTCCCGAGGCCCTGCG
25  CTCTCGGCCGGGCTCCCCAGCCCTCTGCTTTGCCTCCTGCTCCTGGTGCCCCACCACCTC

```

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1:

```

30  M L P G L R R L L Q A P A S A C L L L M L L A L P L A A P S C
   P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S T
   Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N N
   L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E P
   D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L V
   S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F L
35  H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q G
   V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L A
   D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R A
   R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A C
   P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P P
40  A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W G
   G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S A
   G L P S P L L C L L L L V P H H L

```

- 104 -

The following DNA sequence Ngr3 <SEQ ID NO. 3> was identified in mouse:

5 ATGTCTTGGCAGTCTGGAACCACAGTGACACAATCTCCCGTGCAGGCTGCTCAGGTCTCA
GGGTGCTGTGTGGAATTGCTGCTGTTGCTGCTCGCTGGAGAGCTACCTTGGGTGGTGGT
TGTCTCGAGACTGTGTGTGCTACCCTGCGCCCATGACTGTCAGCTGCCAGGCACACAAC
10 TTTGCTGCCATCCCGGAGGGCATCCCAGAGGACAGTGAGCGCATCTTCCTGCAGAACAAAT
CGCATCACCTTCCTCCAGCAGGGCCACTTCAGCCCCGCCATGGTCACCCTCTGGATCTAC
TCCAACAACATCACTTTTCATTGCTCCCAACACCTTCGAGGGCTTTGTGCATCTGGAGGAG
CTAGACCTTGAGACAACCGACAGCTGCGAACGCTGGCACCCGAGACCTTCCAAGGCCTG
15 GTGAAGCTTCACGCCCTCTACCTCTATAAGTGTGGACTGAGCGCCCTGCCCGCAGGCATC
TTTGGTGGCCTGCACAGCCTGCAGTATCTCTACTTGCAGGACAACCATATCGAGTACCTC
CAAGATGACATCTTTGTGGACCTGGTCAATCTCAGTCACTTGTTTCTCCATGGTAACAAG
CTATGGAGCCTGGGCCAAGGCATCTTCCGGGGCCTGGTGAACCTGGACCGGTTGCTGCTG
CATGAGAACCAGCTACAGTGGGTTACACACAAGGCTTTCCATGACCTCCACAGGCTAACC
20 ACCCTCTTTCTCTTCAACAACAGCCTCACTGAGCTGCAGGGTGACTGTCTGGCCCCCCTG
GTGGCCTTGAGAGTTCTTCGCCTCAATGGGAATGCTTGGGACTGTGGCTGCCGGGCACGT
TCCCTGTGGGAATGGCTGCGAAGGTTCCGTGGCTCTAGCTCTGCTGTCCCCTGCGCGACC
CCCGAGCTGCGGCAAGGCCAGGATCTGAAGCTGCTGAGGGTGGAGGACTTCCGGAAGTGC
ACAGGACCAAGTGTCTCTCCTCACCAGATCAAGTCTCACACGCTTACCACCTCTGACAGGGCT
25 GCCCGCAAGGAGCACCATCCGTCCCATGGGGCCTCCAGGGACAAAGGCCACCCACATGGC
CATCCGCCTGGCTCCAGGTCAGGTTACAAGAAGGCAGGCAAGAACTGCACCAGCCACAGG
AACCGBAACCAGATCTCTAAGGTGAGCTCTGGGAAAGAGCTTACCGAACTGCAGGACTAT
GCCCCGACTATCAGCACAAAGTTCAGCTTTGACATCATGCCACCCGACGACCCAAGAGG
AAGGGCAAGTGTGCTCGCAGGACCCCCATCCGTGCCCCAGTGGGGTGCAGCAGGCATCC
TCAGGCACGGCCCTTGGGGCCCCACTCCTGGCCTGGATACTGGGGCTGGCAGTCACTCTC
CGC

The following protein sequence <SEQ ID NO. 4> is deduced protein of SEQ ID
NO:3:

30 MSWQSGTTVTQSPVQAAQVSGCCVELLLLLL
AGELPLGGGCPRDCVCYPAPMTVSCQAHNFA
AIPEGIPEDSERIFLQNNRITFLQQGHFSPA
MVTLWIYSNNITFIAPNTFEGFVHLEELDLG
DNRQLRTLAPETFQGLVKLHALYLYKCLSA
LPAGIFGGLHSLQYLYLQDNHIEYLQDDIFV
35 DLVNLSHLFLHGNKLWSLGQGIFRGLVNLDL
LLHLENQLQWVHHKAFHDLHRLTTLFLFNNS
LTELQGDCLAPLVALEFLRLNGNAWDCGCRA
RSLWEWLRRFRGSSSAVPCATPELRQGQDLK
LLRVEDFRNCTGPVSPHQIKSHTLTTSDDRAA
40 RKEHHPSHGASRDKGHPHGHPPGSRSGYKKA
GKNCTSHRNRNQISKVSSGKELTELQDYAPD
YQHKFSFDIMPTARPKRKGCARRTPIRAPPS
GVQQASSGTALGAPLLAWILGLAVTLR

- 105 -

The following protein sequence <SEQ ID NO. 5> is NgR1 from humans:

MKRASAGGSRL LAWVLWLQAWQVAAPCPGA
 C
 5 CYNEPKVTTSCPQQGLQAVPVGIPAASQRI
 FLHG NRISHVPAASFRA CRNLTLWLHSNVL
 ARIDAAFTGLALLEQLDLSDNAQLRSVDPA
 TFHGLGRLHTLHLDR CGLQELGPGLFRGLAA
 LQYLYLQDNALQALPDDTFRDLGNLTHLFLH
 10 GNRIS SVPERAFRGLHSLDRLL LHQNRVAHV
 HPHA FRDLGRLMTLYLFANNLSALPTEALAP
 LRALQYLR LNDNPWVCD CRARPLWAWLQKFR
 GSSEVPCSLPQRLAGRDLKRLAANDLQGCA
 VATGPYHPIWTGRATDEEPLGLPKCCQPDA
 15 DKASVLEPGRPASAGNALKGRVPPGDSPPGN
 GSGPRHINDSPFGTLPGSAEPPLTAVRPEG
 EPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
 GSGGGGTGDSEGS GALPSLTCSLTPLGLALV
 LWTVLGPC

The following amino acid sequence <SEQ ID NO:6> is a Consensus Sequence of
 20 NgR based on homology with NgR1

C P X X C X C Y X X P X X T X S C X X X X X X X X P X
 X X P X X X X R X F L X X N X I X X X X X X X F X X X X X X X L W X
 X S N X X X X I X X X X F X X X X X L E X L D L X D N X X L R
 25 X X X P X T F X G L X X L X L X L X X C X L X X L X X X X F X
 G L X X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H
 L F L H G N X X X X X X X X X F R G L X X L D R L L L H X N X
 X X X V H X X A F X X L X R L X X L X L F X N X L X X L X X X
 X L A X L X X L X X L R L N X N X W X C X C R A R X L W X W X
 X X X R X S S X V X C X X P X X X X G X D L X X L X X X D X
 30 X X C X X X X P X X P X X X X X X X X X X X X X X X X X X
 X X X X X X X X X X X X X X X X X G X X X X X X X X X X X X
 P P X X X S X
 X
 35 X L X X X X X
 X X X X X L

The following protein sequence <SEQ ID NO:7> is the 66 amino acid active domain
 of Nogo:

40 R I Y K G V I Q A I Q K S D E G H P F R A Y L E S E V A I S E
 E L V Q K Y S N S A L G H V N C T I K E L R R L F L V D D L V
 D S L K

- 106 -

The following protein sequence <SEQ ID NO:8> is the amino acid sequence of the mature NgR2:

5 C P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S
T Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N
N L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E
P D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L
V S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F
L H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q
G V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L
10 A D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R
A R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A
C P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P
P A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W
15 G G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S
A G L P S P L L C L L L L V P H H L

The following protein sequence <SEQ ID NO:9> is the amino acid sequence of the mature NgR3:

20 C P R D C V C Y P A P M T V S C Q A H N F A A I P E G I P E D
S E R I F L Q N N R I T F L Q Q G H F S P A M V T L W I Y S N
N I T F I A P N T F E G F V H L E E L D L G D N R Q L R T L A
P E T F Q G L V K L H A L Y L Y K C G L S A L P A G I F G G L
H S L Q Y L Y L Q D N H I E Y L Q D D I F V D L V N L S H L F
L H G N K L W S L G Q G I F R G L V N L D R L L L H E N Q L Q
25 W V H H K A F H D L H R L T T L F L F N N S L T E L Q G D C L
A P L V A L E F L R L N G N A W D C G C R A R S L W E W L R R
F R G S S S A V P C A T P E L R Q G Q D L K L L R V E D F R N
C T G P V S P H Q I K S H T L T T S D R A A R K E H H P S H G
A S R D K G H P H G H P P G S R S G Y K K A G K N C T S H R N
30 R N Q I S K V S S G K E L T E L Q D Y A P D Y Q H K F S F D I
M P T A R P K R K G K C A R R T P I R A P S G V Q Q A S S G T
A L G A P L L A W I L G L A V T L R

35 The following amino acid sequence <SEQ ID NO:10> is a conserved cysteine motif
(Cysteine domain 1) of the NgR and homologs based on the Consensus Sequence:

CPXXCXC YXXPXXTXSC

- 107 -

The following amino acid sequence <SEQ ID NO:11> is a conserved cysteine motif (Cysteine domain 2) of the NgR and homologs based on the Consensus Sequence:

N X W X C X C R A R X L W X W X X X R X S S S X V X C X X P
X X X X G X D L X X L X X D X X X C

5

The following amino acid sequence <SEQ ID NO:12> is a conserved Leucine-rich domain of the NgR and homologs based on the Consensus Sequence:

R X F L X X N X I X X X X X X F X X X X X X X L W X X S N
X X X X I X X X X F X X X X X L E X L D L X D N X X L R X X X
P X T F X G L X X L X L X L X X C X L X X L X X X X F X G L X
X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H L F L
H G N X X X X X X X X F R G L X X L D R L L H X N X X X X
V H X X A F X X L X R L X X L X L F X N X L X X L X X X X L A
X L X X L X X L R L

10

15

Unless otherwise indicated, X is any amino acid. For example, X where indicated may be no amino acid. Additional features of the invention will be apparent from the following Examples. Examples 1-5 are actual, while the remaining Examples are prophetic.

20

As shown by the following Examples, a gene encoding novel NgRs have been identified by computational analysis of DNA sequence data. The proteins encoded by NgR2 and NgR3 have a putative signal sequence, eight leucine-rich repeat domains in a conserved leucine-rich region (SEQ ID NO:12), a conserved cysteine-rich region (SEQ ID NO:10) N-terminal to the leucine-rich region, a second cysteine-rich domain (SEQ ID NO:11) C-terminal to the leucine-rich region, and a putative glycoposphatidylinositol-linkage (GPI-linkage) site. NgR2 and NgR3 differ from the previously identified NgR sequence. The NgR homologs, when compared to known NgRs, show a consensus sequence (SEQ ID NOs:6). The putative mature NgR2 and NgR3 are shown in Table 5 as SEQ ID NOs: 8 and 9, respectively.

25

30

Example 1: Tblastn query of the HTG database

The protein sequence for the human NgR (NgR1) (SEQ ID NO:5) was used to query the high throughput genomic (HTG) database the use of which is familiar to those skilled in the art. The HTG database is a part of GenBank, a comprehensive

35

- 108 -

NIH genetic sequence database, which includes an annotated collection of all publicly available DNA sequences (*Nucleic Acids Res.* (2000) 28, 15-8). The HTG database includes sequences obtained from genomic DNA. Within genomic DNA, genes are typically encoded by multiple segments of DNA called exons. Thus when one aligns a
5 cDNA sequence (or a protein sequence encoded by a cDNA sequence) to a genomic sequence, the sequence will be broken up into segments depending on the number of exons in the gene.

The BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, (1990) *J. Mol. Biol.* 215,
10 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). The basic BLAST algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some
15 positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased.
20 Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the
25 alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, (1992) *Proc. Natl. Acad. Sci. USA* 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, (1993) *Proc. Natl. Acad. Sci. USA* 90,
30 5873-5787, which is incorporated herein by reference) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which

- 109 -

provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a NgR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a NgR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

To query the HTG database with the NgR protein sequence, we used a variation of the BLAST algorithm known as the tblastn program, which compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames (*J. Mol. Biol.* (1990) 215, 403-410; *Nucleic Acids Res.* (1997) 25, 3389-3402). The results of the tblastn search indicated the presence of genes in the database with a significant identity to the NgR. In addition to finding hits to genomic clones which contain the human and mouse NgR genes, we found hits to clones where the identity was not as high, but still very significant. Three human clones were found (Accession numbers: AC068514, AC016869, AC013606) with an e-value of $4e-43$ and one mouse clone was found (Accession No. AC021768) with an e-value of $1e-78$. The three human clones all appeared to encode the same gene, so further analysis was confined to AC013606.

20 Example 2: Prediction of the human NgR2 protein sequence (AC013606)

The human NgR protein sequence aligned with two regions of translated sequence from nucleotide sequence AC013606, indicating that the new gene was encoded by at least two exons. In order to define the complete gene, we used the computer program GENSCAN[™] (*J. Mol. Biol.* (1997) 268, 78-94) which can identify complete exon/intron structures of genes in genomic DNA. The gene prediction by GENSCAN[™] contained seven exons. By comparing these predicted exons to the NgR, it was concluded that the new human gene contains two of these exons and a part of another (containing the initiating methionine). The predicted cDNA (mRNA) encoded by these three exons was assembled from AC013606 (HTG11; deposited March 2000; length = 143899; GenBank release 118.0; SEQ ID NO: 15) by combining nucleotides from the three exons whose coordinates are: 123292-123322 (exon 1); 130035-130516 (exon 2); and 138589-139335 (exon 3). The sequence for this cDNA

- 110 -

sequence is SEQ ID NO:1 (nucleotide sequence of human NgR2; AC013606). The translation of this cDNA provides the protein sequence of human NgR2 (SEQ ID NO:2).

We used the protein sequence of human NgR2 as a query sequence against the human EST database. A number of hits of high significance were found indicating that the NgR2 mRNA is expressed in a number of tissues including fetal brain. Furthermore, two of these ESTs provided support for the exon structure that we deduced. One EST (Accession No: GB_EST19:AI346757) contains 565 nucleotides corresponding to amino acids 84-271 of the human NgR2 (SEQ ID No:4). This spans the second intron located between amino acids 171 and 172, and provides positive evidence for the splicing of exons 2 and 3 at the mRNA level. Another EST (GB_EST26:AI929019) contains 545 nucleotides, part of which corresponds to amino acids 1-75 of the human NgR2 (SEQ ID NO:2). This spans the first intron located between amino acids 10 and 11, and provides positive evidence for the splicing of exons 1 and 2 at the mRNA level.

Example 3: Prediction of the mouse NgR3 protein sequence (AC021768)

The human NgR protein sequence aligned with only one region of translated sequence from nucleotide sequence AC021768, indicating that most of the new mouse gene was encoded by one large exon. However, upon inspection, the protein encoded by this exon was missing an initiating methionine. In order to define the complete gene, we used the computer program GENSCAN as described above. The gene prediction by GENSCAN contained two exons; the large one found by visual inspection and a short one at the 5' end which provided an initiating methionine. The predicted cDNA (mRNA) encoded by these two exons was assembled from AC021768 (HTG14; deposited March 2000; length = 215980; GenBank release 118.0; SEQ ID NO: 16) by combining nucleotides from the two exons whose coordinates are: the complement of 164265-164325 (exon 1); and the complement of 155671-156992 (exon 2). The sequence for this cDNA sequence is SEQ ID NO:3 (nucleotide sequence of mouse NgR3; AC021768). The translation of this cDNA provides the protein sequence of mouse NgR3 (SEQ ID NO:4).

- 111 -

We used the protein sequence of mouse NgR3 as a query sequence against the mouse EST database. One hit of high significance was found indicating that the NgR2 mRNA is expressed in the heart. This EST (GB_EST20:AI428334) contains 463 nucleotides, part of which correspond to amino acids 45-193 of mouse NgR3 (SEQ ID NO:4).

Example 4: Similarity between the NgRs

An alignment between NgR1 and the two new receptors is shown in Fig. 1A-1B. The similarities between these proteins include:

- 10 (1) The SignalP program, which locates the signal sequence cleavage position, predicts a cleavage before the first conserved cysteine in all the proteins. Thus the mature protein in all cases will have a cysteine at the N-terminus.
- (2) All proteins contain eight Leucine Rich Repeats (LRR). LRRs are short sequence motifs present in a number of proteins with diverse functions and cellular
15 locations. These repeats are usually involved in protein-protein interactions. Each LRR is composed of a beta-alpha unit.
- (3) All three proteins contain a leucine rich repeat N-terminal domain (LRRNT), in which four cysteines are conserved. LRRs are often flanked by cysteine rich domains at both their N and C termini.
- 20 (4) All three proteins contain a LRR C-terminal domain (LRRCT). The LRRCTs of the three NgR proteins can be distinguished from those of other LRR containing proteins, by the pattern of typtophans and cysteines which are completely conserved in this domain.
- (5) All three proteins contain a conserved cysteine in the fourth LRR
25 domain.
- (6) All three proteins contain a conserved potential glycosylation site in the eighth LRR domain.
- (7) NgR2 and NgR3 have a hydrophobic C-terminus, as does NgR1, an indication that they probably also undergo a modification similar to NgR1, where a
30 GPI moiety is covalently linked to a C-terminal amino acid. This allows the protein to remain tethered to the cell.

- 112 -

Example 5: Preparation of Nogo Proteins

A Nogo binding assay was developed which utilizes a method widely used in examining semaphorin and ephrin axonal guidance function (Flanagan & Vanderhaeghen (1998) *Annu. Rev. Neurosci.* 21,3 09-345; Takahashi *et al.*, (1999) *Cell* 99, 59-69). It involves fusing a secreted placental alkaline phosphatase (AP) moiety to the ligand in question to provide a biologically active receptor binding agent which can be detected with an extremely sensitive colorimetric assay. For Nogo, an expression vector is created encoding a signal peptide, a His6 tag for purification, AP, and the 66 amino acid active domain of Nogo. The fusion protein can be purified from the conditioned medium of transfected cells in milligram amounts. This protein is biologically active as a growth cone collapsing agent with an EC₅₀ of 1 nM.

Alternatively, a glutathione-S-transferase Nogo (GST-Nogo) fusion protein may be prepared. For GST-Nogo, an expression vector (*e.g.*, a pGEX vector) is created encoding a signal peptide, GST, and the 66 amino acid active domain of Nogo. GST-Nogo may be purified from the culture medium and used as a GST fusion protein, or GST may be cleaved from the Nogo portion of the fusion protein with an enzyme that recognizes the specific amino acid cleavage site engineered between the GST portion and the Nogo portion of the fusion protein. Such sites are part of the commercially available GST vectors. The specific cleavage sites and enzymes may be used in accordance with the Manufacturer's specifications.

It has been found that AP-Nogo is actually slightly more potent than GST-Nogo, perhaps because the protein is synthesized in a eukaryotic rather than a prokaryotic cell.

Binding of Nogo to immobilized NgR homologs may be performed in an ELISA-type assay in which AP-Nogo is allowed to react with an immobilized receptor homolog. Specificity of binding may be demonstrated in a competitive binding assay using increasing amounts of GST-Nogo in the type of assay to show a decreasing amount of binding of AP-Nogo (as judged in the colorimetric assay).

Example 6: Transfected COS Cell binding Assays

The homologs of the present invention may be used in transfection studies in COS cells to demonstrate binding of Nogo. Specifically, nucleotide sequences

- 113 -

encoding NgR2 and NgR3 may be transfected into COS cells using a suitable vector. Non-transfected COS-7 cells do not bind AP-Nogo. However, transfection of COS cells with nucleic acid sequences encoding NgRs will make them capable of binding Nogo. AP alone does not bind with any stable affinity to these transfected cells,
5 indicating that any affinity of Nogo for NgR2 or NgR3 would be due to the 66 amino acids derived from Nogo. Furthermore, specific affinity of Nogo for the NgR2 or NgR3 proteins may be tested in displacement of AP-Nogo assays using GST-Nogo. NgR2 and/or NgR3 may also bind homologs of Nogo, which may also be tested using this assay.

10

Example 7: Expression of NgR in Human Cell Lines using Northern Blot and a Random-Primed Probe

A Northern blot is purchased from a commercial source, or RNA samples from cells of interest are run on an agarose gel and blotted to a membrane using any of the well known techniques for Northern blotting. The blot is probed with a fragment of
15 NgR2 (SEQ ID NO:1) or NgR3 (SEQ ID NO:3). The probe is prepared from 50 ng of cDNA labeled by a random-primed method (Feinberg and Vogelstein (1983) *Anal. Biochem.* 132, 6-13). Hybridization is carried out at 68°C for 1 hour in ExpressHyb™ solution (Clontech, Cat. No. 8015-1) followed by washing with 2X
20 SSC/0.05% SDS at room temperature and two washes with 0.1X SSC/0.1% SDS at 50°C. Expression of NgR2 and/or NgR3 can be assessed by the presence of an appropriately sized band on the blot.

25

Example 8: Cloning of cDNA corresponding to NgRs

To obtain the full-length clone corresponding to NgR2 from a cDNA library, the following method may be used. A cDNA library is generated using standard methods from a tissue known to contain NgR2. Such a tissue was identified in Example 2. 1×10^6 plaque forming units from the cDNA library may be screened in
30 duplicate on OPTITRAN™ filters. The filters are hybridized with ^{32}P -labeled oligonucleotides that are generated from the ESTs corresponding to portions of NgR2. The hybridization reaction may consist of 400 mls plaque screen buffer (50mM Tris pH

- 114 -

7.5, 1M NaCl, 0.1% Sodium pyrophosphate, 0.2% Polyvinylpyrrolidone and 0.2% Ficoll) containing 10% Dextran sulfate and 100µg/ml tRNA and 80 pmol each ³²P-labeled oligonucleotide at 65°C overnight. The filters are washed twice with 2X SSC/1%SDS and twice with 1X SSC/1%SDS and exposed to film. Duplicate positives
5 are purified. DNA from each of these clones is analyzed by restriction enzyme digest followed by agarose gel electrophoresis and Southern blotting. The filters are hybridized to the ³²P-labeled oligonucleotides used for the original hybridization to confirm that inserts hybridize to the probe. The insert is then sequenced to confirm that it represents the cDNA for NgR2. Similar methods may be used to generate a
10 full-length clone corresponding to NgR3.

Alternatively, a full-length clone of NgR2 or NgR3 can be obtained by a person of ordinary skill in the art employing conventional PCR techniques.

Example 9: Hybridization Analysis to demonstrate NgR expression in the brain

15 The expression of NgR in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and sagittal rat brain cryosections (20 µm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are
20 processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at
25 room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A NgR-specific probe may be generated using PCR. Following PCR amplification, the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for
30 the sense strand of NgR, a cloned NgR fragment cloned in pBluescript II may be linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (*i.e.*, cRNA riboprobes) using the vector-borne T7 promoter and

- 115 -

commercially available T7 RNA polymerase. A probe specific for the antisense strand of NgR may also be readily prepared using the NgR clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3 promoter and cognate polymerase. The riboprobes may be labeled with [³⁵S]-UTP to yield a specific activity of about 0.40×10^6 cpm/pmol for antisense riboprobes and about 0.65×10^6 cpm/pmol for sense-strand riboprobes. Each riboprobe may be subsequently denatured and added (2 pmol/ml) to hybridization buffer which contains 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol. Microscope slides containing sequential brain cryosections may be independently exposed to 45 µl of hybridization solution per slide and silanized cover slips may be placed over the sections being exposed to hybridization solution. Sections are incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are then exposed to sense or antisense NgR-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment involving the exposure of slides to 20 µg/ml RNase A in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three high-stringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMax™ MR-1 film. After 13 days of exposure, the film is developed, and any significant hybridization signal is detected. Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with Kodak NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violet-stained cell bodies. Autoradiographic grains found between cell bodies indicate non-specific binding of the probe.

- 116 -

In some cases, such as using a probe to detect a NgR homolog in a heterologous species, in order to achieve optimal hybridization, it may be necessary to decrease the stringency conditions. Such conditions are well known to those of ordinary skill in the art and examples are provided above.

5 Expression of NgR in the brain provides an indication that modulators of NgR activity have utility for treating neurological disorders. Some other diseases for which modulators of NgR may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of NgR modulators, including NgR ligands and anti-NgR antibodies, to treat individuals having
10 such disease states is intended as an aspect of the invention.

Example 10: Northern Blot Analysis of NgR-RNA with a PCR-generated Probe

Northern blot hybridizations may be performed to examine the expression of NgR mRNA. A clone containing at least a portion of the sequence of SEQ ID NO:1
15 may be used as a probe. Vector-specific primers are used in PCR to generate a hybridization probe fragment for ³²P-labeling. The PCR is performed as follows:

	Mix:	1μl	NgR-containing plasmid
		2μl	fwd primer (10-50 pM)
20		2μl	rev primer (10-50 pM)
		10μl	10xPCR buffer (such as that provided with the enzyme, Amersham Pharmacia Biotech)
		1μl	10mM dNTP (such as #1 969 064 from Boehringer Mannheim)
		0.5μl	Taq polymerase (such as #27-0799-62, Amersham Pharmacia Biotech)
25		83.5μl	water

- 117 -

PCR is performed in a Thermocycler using the following program:

	94°C	5min		
5	94°C	1min		30 cycles
	55°C	1min		
	72°C	1min		
10	72°C	10min		

The PCR product may be purified using QIAquick PCR Purification Kit (#28104) from Qiagen, and radiactively labeled with ^{32}P -dCTP (#AA0005/250, Amersham Pharmacia Biotech)) may be done by random priming using "Ready-to-go DNA Labeling Beads" (#27-9240-01) from Amersham Pharmacia Biotech.

Hybridization is carried out on Human Multiple Tissue Northern Blot from Clontech as described in manufacturer's protocol, or on a Northern Blot prepared by running RNA samples from cells of interest on an agarose gel and blotting to a membrane using any of the known Northern blotting protocols. After exposure overnight on Molecular Dynamics Phosphor Imager screen (#MD146-814) bands of an appropriate size are visualized.

Example 11: Recombinant Expression of NgR in Eukaryotic Host Cells

A. Expression of NgR in Mammalian Cells

To produce NgR protein, a NgR-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, a NgR-encoding sequence described in Table 4 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells using the transfection reagent FuGENE6™ (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing NgR are selected by growth in the presence of 100 µg/ml zeocin (Stratagene, LaJolla, CA). As an alternative to FuGENE6™, the expression vector may carry the gene for dihydrofolate reductase (dhfr) and selection of clones with methotrexate (MTX) drug pressure

- 118 -

allows for stable transformation of CHO cells. Optionally, NgR may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the NgR amino acid sequence, and the antisera is used to affinity purify Nogo-R.

- 5 The NgR also may be expressed in-frame with a tag sequence (*e.g.*, polyhistidine, hemagglutinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for NgR polypeptides, such as assays described below, do not require purification of NgR from the host cell.

10 **B. Expression of NgR in CHO cells**

- For expression of NgR in Chinese hamster ovary (CHO) cells, a plasmid bearing the relevant NgR coding sequence is prepared, using a vector which also bears the selectable marker dihydrofolate reductase (DHFR). The plasmid is transfected into CHO cells. Selection under MTX drug pressure allows for preparation of stable
- 15 transformants of a NgR (NgR2 or NgR3) in an expression plasmid carrying a selectable marker such as DHFR.

C. Expression of NgR in 293 cells

- For expression of NgR in mammalian cells 293 (transformed human, primary
- 20 embryonic kidney cells), a plasmid bearing the relevant NgR coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for
- 25 selection of stable transfectants. The forward primer for amplification of this NgR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the HindIII cloning site and nucleotides matching the NgR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning
- 30 and nucleotides corresponding to the reverse complement of the NgR sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII*-*XhoI* sites of the vector.

- 119 -

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAP™ transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-NgR peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by Western blots probed with anti-His, anti-Myc or anti-NgR peptide antibodies.

D. Transient Expression of Nogo-R in COS cells

For expression of the NgR in COS7 cells, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, for example, can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an *Xba*I restriction site for cloning, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1. The reverse primer is also determined by routine procedures and preferably contains 5'-extension of nucleotides which introduces a *Sa*I cloning site followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1.

The PCR consists of an initial denaturation step of 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the *Xba*I and *Sa*I sites of vector p3-CI. This construct is transformed into *E. coli* cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

NgR expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for

- 120 -

example, chromatography. Purified NgR is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C. NgR3 may also be expressed using this method and the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:13.

5

E. Expression of NgR in Insect Cells

For expression of NgR in a baculovirus system, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO:1, 3 or 13 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively). The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13, respectively).

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (Pharmingen, San Diego, CA). The pACHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pACHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of NgR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (1987) A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555.

- 121 -

In a preferred embodiment, pAcHLT-A containing NgR gene is introduced into baculovirus using the "BaculoGold™" transfection kit (Pharmingen, San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with ³⁵S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a NgR polypeptide in a Sf9 cells, a polynucleotide molecule having the nucleotide sequence of SEQ ID NO:1 (or SEQ ID NO:3 or SEQ ID NO:13) can be amplified by PCR using the primers and methods described above for baculovirus expression. The NgR cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the NgR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

F. Expression of soluble forms of NgR2 and NgR3 as NgR-Ig fusion proteins.

To generate a NgR2-Ig fusion protein, standard methods may be used as described in the literature (*e.g.* Sanicola *et al.* (1997) *Proc. Natl. Acad. Sci. USA*. 94, 6238-6243). For example, a DNA fragment encoding NgR2 without the sequence encoding the hydrophobic C-terminus (GPI anchor signal) may be ligated to a DNA fragment encoding the Fc domain of IgG1 (which may be human IgG1), and the chimeric fragment may be cloned into an expression vector to generate a plasmid. The plasmid may then be transfected into Chinese hamster ovary cells to generate a stable cell line producing the fusion protein. The fusion protein is then purified from conditioned media using standard methods. For example, clarified conditioned media from the cell line may be loaded by gravity directly onto Protein A Sepharose. The column may then be washed with five column volumes each of PBS, PBS containing

- 122 -

0.5 M NaCl, and 25 mM sodium phosphate, 100 mM NaCl (pH 5.0). The bound protein may then be eluted with 25 mM NaH₂PO₄, 100 mM NaCl (pH 2.8) and immediately neutralized with 1/10 fraction volume of 0.5 M Na₂HPO₄ (pH 8.6).

Similar methods may be used to generate a NgR3-Ig fusion protein.

5

Example 12: Interaction Trap/Two-Hybrid System

In order to assay for NgR-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.* (1989) *Nature* 340, 245, which is incorporated herein by reference in its entirety. A protocol is published in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 1999, John Wiley & Sons, NY and Ausubel, F. M. *et al.* 1992, SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

15 A fusion of the nucleotide sequences encoding all or partial NgR and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (*i.e.*, pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7) from cDNA of potential NgR-binding proteins (for protocols on forming cDNA libraries, see Sambrook *et al.* 1989, MOLECULAR CLONING: A LABORATORY MANUAL, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/NgR fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of transcriptional activity. Yeast cells are transformed (ca. 10⁵ transformants/mg DNA) with both the NgR and library fusion plasmids according to standard procedure (Ausubel, *et al.*, 1992, SHORT PROTOCOLS IN MOLECULAR BIOLOGY, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety). *In vivo* binding of DNA-BD/NgR with AD/library proteins results in transcription of specific yeast plasmid reporter genes (*i.e.*, *lacZ*, *HIS3*, *ADE2*, *LEU2*). Yeast cells are plated on nutrient-deficient

- 123 -

media to screen for expression of reporter genes. Colonies are dually assayed for β -galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl-b-D-galactoside) supplemented media (filter assay for β -galactosidase activity is described in Breeden *et al.*, (1985) *Cold Spring Harb. Symp. Quant. Biol.*, 50, 643, which is incorporated herein by reference in its entirety).

5 Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific NgR/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine

10 the identity of the NgR-binding protein.

Example 13: Antibodies to Nogo-R

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the NgR receptor, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be

15 found, for example, in Sambrook *et al.* (1989), above, and Harlow *et al.* (Eds.), ANTIBODIES A LABORATORY MANUAL; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988). In one embodiment, recombinant NgR polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate

20 the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of NgR (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of Nogo-R, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a

25 hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant NgR or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger

30 mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Limpet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with

- 124 -

Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of NgR antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to confirm the presence of antibodies that immunoreact with NgR. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize NgR. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, UT) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 µM sodium hypoxanthine, 0.4 µM aminopterin, 16 µM thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5×10^6 thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning, NY).

On days 2, 4, and 6 after the fusion, 100 µl of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to NgR. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-NgR antibodies are obtained.

- 125 -

B. Humanization of anti-NgR monoclonal antibodies

The expression pattern of NgR as reported herein and the potential of NgRs as targets for therapeutic intervention suggest therapeutic indications for NgR inhibitors (antagonists). NgR-neutralizing antibodies comprise one class of therapeutics useful as
5 NgR antagonists. Following are protocols to improve the utility of anti-NgR monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-NgR antibodies).

The principles of humanization have been described in the literature and are
10 facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest
15 with the constant domains of human antibody molecules. (See, *e.g.*, Morrison *et al.*, (1989) *Adv. Immunol.*, 44, 65-92). The variable domains of NgR-neutralizing anti-NgR antibodies are cloned from the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and
20 the resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned
25 into human antibody sequences. (See, *e.g.*, Jones *et al.*, (1986) *Nature* 321, 522-525; Riechmann *et al.*, (1988) *Nature* 332, 323-327; Verhoeyen *et al.*, (1988) *Science* 239, 1534-1536 ; and Tempest *et al.*, (1991) *Bio/Technology* 9, 266-271). If necessary, the β -sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding
30 domain of the original monoclonal antibody. (See Kettleborough *et al.*, (1991) *Protein Engin.* 4, 773-783; and Foote *et al.*, (1992) *J. Mol. Biol.* 224, 487-499).

- 126 -

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan (1991) *Mol. Immunol.* 28, 489-498.

5 The foregoing approaches are employed using NgR-neutralizing anti-NgR monoclonal antibodies and the hybridomas that produce them to generate humanized NgR-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein NgR expression or ligand-mediated NgR signaling is detrimental.

10 **C. Human NgR-Neutralizing Antibodies from Phage Display**

Human NgR-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.* (1997) *Human Antibodies* 8, 155-168; Hoogenboom (1997) *TIBTECH* 15, 62-70; and Rader *et al.* (1997), *Curr. Opin. Biotechnol.* 8, 503-508, all of which are incorporated by reference. For example,
15 antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is
20 expressed in bacteria, and the library is screened for NgR-specific phage-antibodies using labeled or immobilized NgR as antigen-probe.

D. Human NgR-neutralizing antibodies from transgenic mice

Human NgR-neutralizing antibodies are generated in transgenic mice essentially
25 as described in Bruggemann *et al.* (1996) *Immunol. Today* 17, 391-397 and Bruggemann *et al.* (1997) *Curr. Opin. Biotechnol.* 8, 455-458. Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a NgR composition using conventional immunization protocols. hybridomas are generated using B cells from the
30 immunized mice using conventional protocols and screened to identify hybridomas secreting anti-NgR human antibodies (*e.g.*, as described above).

- 127 -

Example 14: Assays to Identify Modulators of NgR Activity

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of NgR activity. Among the modulators that can be identified by these assays are natural ligand compounds of the receptor; synthetic
5 analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind NgR are useful for identifying NgR in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes,
10 and the like). Agonist and antagonist modulators are useful for up-regulating and down-regulating NgR activity, respectively, to treat disease states characterized by abnormal levels of NgR activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or *visa versa*).

15

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in NgR-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See,
20 *e.g.*, Sutherland *et al.*, (1968) *Circulation* 37, 279; Frandsen *et al.*, (1976) *Life Sciences* 18, 529-541; Dooley *et al.*, (1997) *J. Pharmacol. Exp. Therap.* 283, 735-41; and George *et al.*, (1997) *J. Biomol. Screening* 2, 235-40). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

25 Briefly, the NgR coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well
30 microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with

- 128 -

wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (*i.e.*, candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [¹²⁵I]-labeled cAMP, and the plate is counted using a Packard TopcountTM 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) and fixed amounts of [¹²⁵I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of NgR modulating activity. Modulators that act as agonists of receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP levels. Agonists of receptors which couple to the G_{i/o} subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

20 **B. Aequorin Assays**

In another assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a NgR expression construct and a construct that encodes the photoprotein apoaequorin. In the presence of the cofactor coelenterazine, apoaequorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. (See generally, Cobbold, *et al.* "Aequorin measurements of cytoplasmic free calcium," *In*: McCormack J.G. and Cobbold P.H., eds., CELLULAR CALCIUM: A PRACTICAL APPROACH. Oxford: IRL Press (1991); Stables *et al.*, (1997) *Anal. Biochem.* 252, 115-26; and Haugland, HANDBOOK OF FLUORESCENT PROBES AND RESEARCH CHEMICALS. Sixth edition. Molecular Probes, Eugene, OR (1996)).

30 In one exemplary assay, NgR is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaequorin (Molecular Probes, Eugene, OR) into CHO cells

- 129 -

using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to serum-free MEM containing 5 µM coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate NgR modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 µl/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 µl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds.

Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC₅₀ values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for assaying for modulators of NgR activity. Cells (*e.g.*, CHO cells or COS 7 cells) are transiently co-transfected with both a NgR expression construct (*e.g.*, NgR in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George *et al.* (1997) *J. Biomol. Screening* 2, 235-240; and Stratowa *et al.* (1995) *Curr. Opin. Biotechnol.* 6, 574-581). Luciferase

- 130 -

activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in
5 MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with both a NgR expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (Legally, CA). Transfections are
10 performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are
15 incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450
20 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3-20-fold stimulation of
25 luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of
30 receptor activity, and such assays can be employed to screen for modulators of NgR activity. For example, CHO cells stably transfected with a NgR expression vector are plated at a density of 4×10^4 cells/well in Packard black-walled, 96-well plates

- 131 -

5 specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3™
5 AM, Fluo-4™ AM, Calcium Green™-1 AM, or Oregon Green™ 488 BAPTA-1 AM), each at a concentration of 4 μM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of
10 the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μM; positive control), or ATP (4 μM; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.*
15 (1998) *Drug Dev. Res.* 44,14-20). The F-stop for the detector camera is set at 2.5 and the length of exposure is 0.4 milliseconds. Basal fluorescence of cells is measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level is subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium
20 ionophore A23187 and ATP increase the calcium signal 200% above baseline levels. In general, activated NgRs increase the calcium signal at least about 10-15% above baseline signal.

E. [³⁵S]GTPγS Binding Assay

25 It is also possible to evaluate whether NgR signals through a G protein-mediated pathway. Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [³⁵S]-GTPγS in the presence and absence of candidate modulators provides another assay
30 for modulator activity. (See, *e.g.*, Kowal *et al.*, (1998) *Neuropharmacology* 37, 179-187.).

- 132 -

In one exemplary assay, cells stably transfected with a NgR expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE
5 buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (1 ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to
10 isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM MgCl_2 , 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C
15 are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM MgCl_2 , 1 mM EDTA, 120 mM NaCl, 10 μM GDP, and 0.2 mM ascorbate, at a concentration of 10-50 $\mu\text{g}/\text{ml}$. In a final volume of 90 μl , homogenates are incubated with varying concentrations of candidate modulator compounds or 100 μM GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 μl guanosine
20 $5'\text{-O-(3[}^{35}\text{S]thio) triphosphate}$ (NEN, 1200 Ci/mmol; $[\text{}^{35}\text{S}]\text{-GTP}\gamma\text{S}$), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl_2 , at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with
25 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM MgCl_2 . Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of $[\text{}^{35}\text{S}]\text{-GTP}\gamma\text{S}$ is measured in the presence of 100 μM GTP and subtracted from the total. Compounds are selected that modulate the amount of $[\text{}^{35}\text{S}]\text{-GTP}\gamma\text{S}$ binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in
30 $[\text{}^{35}\text{S}]\text{-GTP}\gamma\text{S}$ binding. This response is blocked by antagonists.

F. $[\text{}^3\text{H}]\text{Arachidonic Acid Release}$

- 133 -

The activation of NgRs may also potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of NgR activity. (See, *e.g.*, Kanterman *et al.*, (1991) *Mol. Pharmacol.* 39,364-369.) For example, CHO cells that are stably transfected with a NgR expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [³H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 µCi/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10 µM ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [³H]-arachidonic acid. This potentiation is blocked by antagonists.

G. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of NgR activity are assayed by monitoring extracellular changes in pH induced by the test compounds (see, *e.g.*, Dunlop *et al.* (1998) *J. Pharmacol. Toxicol. Meth.* 40, 47-55). In one embodiment, CHO cells transfected with a NgR expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4×10^5 cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 µg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO₂ for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 µg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 µl/minute.

- 134 -

Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

Example 15: mNgR3 does not bind hNogo-A(1055-1120)

To functionally test the mouse NgR3 (hereinafter, mNgR3) for its ability to bind hNogo-A(1055-1120), a cDNA expression vector for a myc epitope-tagged mNgR3 protein was created. The mouse NgR3 cDNA was amplified by PCR from mouse adult brain cDNA, from the signal sequence to the stop codon, and ligated into the pSecTag2 vector such that the vector encodes a signal sequence followed by a myc tag followed by the mature mNgR3 sequence. This plasmid was transfected into COS07 cells, and expression of a myc-tagged protein of the predicted size was verified by immunoblot analysis. Alkaline phosphatase-hNogo-A(1055-1120) binding studies and myc immunohistology were conducted as described (Fournier et al., supra).

The cells expressing mNgR3 express the myc-tagged protein but binding to AP-hNogo-A(1055-1120) was not observed under the conditions employed (Fig. 8).

Example 16: Identification of partial human NgR3 cDNA and protein sequences

The tblastn program was used to search for the human homolog of mouse NgR3. The mouse NgR3 protein sequence (SEQ ID NO:4) was used to query a proprietary human expressed sequence tag (EST) database from Incyte yielding one highly significant hit: Incyte Template ID 190989.1. This sequence (937 nucleotides)

- 135 -

contains an open reading frame of 312 amino acids in the second reverse frame that exhibits 88% identity with residues 66 to 381 of mouse NgR3 (SEQ ID NO:4), strongly indicating that it is part of the human NgR3 homolog.

5 A query of SEQ ID NO:4 against the public human EST database in Genbank also produced a hit with a 465-bp EST (Accession number: R35699; Version number: R35699.1; GI: 792600). There are a number of single nucleotide deletions and insertions within this sequence which cause frame shift errors. All of the reliable sequence contained in this public EST is present in the Incyte EST (Template ID 190989.1).

10 To obtain more nucleotide sequence that would extend the amino acid sequence at that carboxy terminal end, the I.M.A.G.E. Consortium clone No. 38319, which corresponds to Genbank accession No. R35699, was purchased from Incyte Genomics Inc. and subjected to further DNA sequence analysis. This clone consists of a NotI/HinD III fragment containing the sequence of interest, cloned into the
15 NotI/HinD III sites of the vector Lfamid BA (<http://image.llnl.gov/image/html/libs/lafmidBA.shtml>). The clone was received as an agar stab, which was streaked out on LB agar plates containing 50ug/ml ampicillin to isolate individual colonies. Six colonies were grown in LB medium with antibiotic, and plasmid DNA was prepared using the Promega Wizard Plus Miniprep DNA
20 Purification System (Promega #A7500). These DNAs were subsequently digested with NotI and HinD III restriction enzymes to confirm that the clones contained an insert. The insert of one isolate was sequenced using a combination of vector specific and gene specific primers yielding a partial nucleotide sequence of human NgR3 of 1176 nucleotides (SEQ ID NO:13). A translation of this sequence provides a partial
25 sequence for human NgR3 of 392 amino acids (SEQ ID NO:14).

The nucleotide sequence of SEQ ID NO:13 differs from the Incyte EST sequence at three positions. Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in the Incyte Template ID 190989.1 are GT (i.e., positions 12-13 of the complement of Incyte Template ID 190989.1). In
30 addition, position 641 in SEQ ID NO:13 is a C, whereas the corresponding nucleotide in the Incyte Template ID 190989.1 sequence is an A (i.e., position 641 of the complement of Incyte Template ID 190989.1). This results in two changes in amino

- 136 -

acids when comparing SEQ ID NO:14 to the ORF encoded by Incyte Template 190989.1: SEQ ID NO:14 contains a valine at position 5, whereas the ORF encoded by Incyte Template ID 190989.1 contains a leucine; SEQ ID NO:14 contains an alanine at position 214, whereas the ORF encoded by Incyte Template ID 190989.1
5 contains a glutamic acid.

The nucleotide sequence of SEQ ID NO:13 differs from the public EST (Accession number: R35699; Version number: R35699.1; GI: 792600) sequence at two positions (within the first 200 nucleotides of reliable sequence). Nucleotide positions 12-13 in SEQ ID NO:13 are CG, whereas the corresponding nucleotides in
10 the public EST are GT (i.e., positions 12-13 of the public EST; Accession no: R35699; Version no: R35699.1; GI: 792600) This leads to a single amino acid change when comparing SEQ ID NO:14 to the ORF encoded by the public EST: SEQ ID NO:14 contains a valine at position 5, while the ORF encoded by the public EST contains a leucine.

15 A Bestfit analysis of the partial human amino acid sequence with the full-length mouse amino acid sequence indicates that the human NgR3 amino acid sequence is complete at the carboxy terminal end and that they share 89.54% identity. An alignment of all the NgR proteins is shown in Figure 9. Although the human NgR3 amino acid sequence is missing the first 25 amino acids, it can be determined that the
20 human NgR3 protein contains the following features in common with the other NgR sequences: (1) eight Leucine Rich Repeat (LRR) domains; (2) an LRR carboxy-terminal (LRR-CT) domain; (3) a conserved cysteine in the fourth LRR domain; (4) a conserved potential glycosylation site in the eighth LRR domain; and (5) a hydrophobic carboxyl terminus.

25 As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by
30 reference. This application claims benefit from United States provisional application 60/238,361, filed October 6, 2000, which is incorporated by reference herein in its entirety.

- 137 -

Key for Sequence Listing:

	SEQ ID NO:1	human NgR2 cDNA sequence derived from genomic sequence AC013606
	SEQ ID NO:2	human NgR2 amino acid sequence
5	SEQ ID NO:3	mouse NgR3 cDNA sequence derived from AC021768
	SEQ ID NO:4	a mouse NgR3 amino acid sequence
	SEQ ID NO:5	a human NgR1 amino acid sequence
	SEQ ID NO:6	a consensus amino acid sequence for NgRs
	SEQ ID NO:7	#1055-1120 amino acid residues of hNogoA (Nogo-66)
10	SEQ ID NO:8	a mature human NgR2 amino acid sequence
	SEQ ID NO:9	a mature mouse NgR3 amino acid sequence
	SEQ ID NO:10	a consensus NgR LLRNT amino acid sequence
	SEQ ID NO:11	a consensus NgR LRRCT domain amino acid sequence
	SEQ ID NO:12	a consensus NgR LRR domain amino acid sequence
15	SEQ ID NO:13	a partial human NgR3 nucleotide sequence
	SEQ ID NO:14	a partial human NgR3 amino acid sequence
	SEQ ID NO:15	a genomic sequence encoding a human NgR2 sequence.
	SEQ ID NO:16	a genomic sequence (complementary strand) encoding a mouse NgR3
20	SEQ ID NO:17	a mouse NgR1 amino acid sequence
	SEQ ID NO:18	a consensus sequence for the NTLRRCT domain of NgR
	SEQ ID NO:19	an consensus NgR LRRCT domain amino acid sequence

25

30

- 138 -

CLAIMS

What is claimed is:

5

1. An isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide comprising an LRRCT domain consisting of the amino acid sequence:

10

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

15

wherein X is any amino acid or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

20

2. The isolated nucleic acid according to claim 1, wherein X₁₇ and X₂₃ are independently selected from the group consisting of: arginine and lysine.

25

3. The isolated nucleic acid according to claim 2, wherein the amino acid sequence of the LRRCT domain is selected from the group consisting of: residues #261-310 of SEQ ID NO:2 and residues 261-310 of SEQ ID NO: 2 with up to 10 conservative amino acid substitutions.

30

4. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 2.

5. An isolated nucleic acid encoding the polypeptide of SEQ ID NO: 4 (mouse NgR3) or SEQ ID NO: 14 (human NgR3).

6. The isolated nucleic acid according to claim 1, wherein the

- 139 -

polypeptide comprises: (a) a NTLRRCT domain, and (b) less than a complete CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

5 7. The isolated nucleic acid to claim 1, wherein the polypeptide does not comprise an intact GPI domain.

8. An isolated nucleic acid consisting essentially of a nucleotide sequence complementary to a nucleotide sequence encoding a polypeptide selected
10 from the group consisting of: a polypeptide consisting of residues 311-395 of SEQ ID NO: 2, a polypeptide consisting of residues 256-396 of SEQ ID NO: 14 and a polypeptide consisting of residues 321-438 of SEQ ID NO: 4, wherein the nucleic acid is from 8 to 100 nucleotides in length.

15 9. A vector comprising the nucleic acid of any one of claims 1, 4 or 5.

10. A host cell comprising a vector according to claim 9.

11. A polypeptide comprising a LRRCT amino acid sequence:

20

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

25

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

wherein X is any amino acid residue or a gap and the polypeptide does not comprise the amino acid sequence from residue 260 to 309 of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17 (mouse NgR1).

30

12. The polypeptide according to claim 11, wherein X₁₇ and X₂₃ is selected from the group consisting of arginine and lysine.

- 140 -

13. The polypeptide according to claim 11, wherein X₁₉ is glycine.

[SEQ ID NO:11]

14. The polypeptide according to claim 11, wherein the amino acid
5 sequence is selected from the group consisting of residues 261–310 of SEQ ID NO:2, residues 206–255 of SEQ ID NO: 14, residues 271–320 of SEQ ID NO:4 and amino acid sequences thereof comprising a conservative substitution.

15. A polypeptide comprising a NTLRRCT amino acid sequence:

10

15

20

C P X₁ X₂ C X₃ C Y X₄ X₅ P X₆ X₇ T X₈ S C X₉ X₁₀ X₁₁ X₁₂ X₁₃ X₁₄ X₁₅ X₁₆ P
X₁₇ X₁₈ X₁₉ P X₂₀ X₂₁ X₂₂ X₂₃ R X₂₄ F L X₂₅ X₂₆ N X₂₇ I X₂₈ X₂₉ X₃₀ X₃₁ X₃₂ X₃₃
X₃₄ F X₃₅ X₃₆ X₃₇ X₃₈ X₃₉ X₄₀ X₄₁ X₄₂ L W X₄₃ X₄₄ S N X₄₅ X₄₆ X₄₇ X₄₈ I X₄₉
X₅₀ X₅₁ X₅₂ F X₅₃ X₅₄ X₅₅ X₅₆ X₅₇ L E X₅₈ L D L X₅₉ D N X₆₀ X₆₁ L X₆₂ X₆₃ X₆₄
X₆₅ P X₆₆ T F X₆₇ G L X₆₈ X₆₉ L X₇₀ X₇₁ L X₇₂ L X₇₃ X₇₄ C X₇₅ L X₇₆ X₇₇ L X₇₈
X₇₉ X₈₀ X₈₁ F X₈₂ G L X₈₃ X₈₄ L Q Y L Y L Q X₈₅ N X₈₆ X₈₇ X₈₈ X₈₉ L X₉₀ D
X₉₁ X₉₂ F X₉₃ D L X₉₄ N L X₉₅ H L F L H G N X₉₆ X₉₇ X₉₈ X₉₉ X₁₀₀ X₁₀₁ X₁₀₂
X₁₀₃ X₁₀₄ F R G L X₁₀₅ X₁₀₆ L D R L L L H X₁₀₇ N X₁₀₈ X₁₀₉ X₁₁₀ X₁₁₁ V H X₁₁₂
X₁₁₃ A F X₁₁₄ X₁₁₅ L X₁₁₆ R L X₁₁₇ X₁₁₈ L X₁₁₉ L F X₁₂₀ N X₁₂₁ L X₁₂₂ X₁₂₃ L
X₁₂₄ X₁₂₅ X₁₂₆ X₁₂₇ L X₁₂₈ X₁₂₉ L X₁₃₀ X₁₃₁ L X₁₃₂ X₁₃₃ L R L N X₁₃₄ N X₁₃₅ W
X₁₃₆ C X₁₃₇ C R X₁₃₈ R X₁₃₉ L W X₁₄₀ W X₁₄₁ X₁₄₂ X₁₄₃ X₁₄₄ R X₁₄₅ S S S X₁₄₆
V X₁₄₇ C X₁₄₈ X₁₄₉ P X₁₅₀ X₁₅₁ X₁₅₂ X₁₅₃ X₁₅₄ X₁₅₅ D L X₁₅₆ X₁₅₇ L X₁₅₈ X₁₅₉ X₁₆₀
D X₁₆₁ X₁₆₂ X₁₆₃ C [SEQ ID NO:18]

25

wherein X is any amino acid residue or a gap and wherein the polypeptide is
not the polypeptide of SEQ ID NO: 5 (human NgR1) or SEQ ID NO: 17
(mouse NgR1).

30

16. The polypeptide according to claim 15, wherein X₆, X₃₇ and X₃₈
represents a gap.

17. A polypeptide comprising an amino sequence selected from the

- 141 -

group consisting of: SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:14.

18. The polypeptide according any one of claims 11, 15 or 17, wherein the polypeptide comprises: (a) an NTLRRCT domain, and (b) less than a complete
5 CTS domain, provided that a partial CTS domain, if present, consists of no more than the first 39 amino acids of the CTS domain.

19. The polypeptide according to any one of claims 11, 15 or 17, wherein the polypeptide does not comprise an intact GPI domain.
10

20. The polypeptide according to any one of claims 11, 15 or 17, wherein the amino acid sequence of the polypeptide further comprises an amino acid sequence of a heterologous polypeptide.

21. The polypeptide according to claim 20, wherein the heterologous polypeptide is an Fc portion of an antibody.
15

22. A method of producing a polypeptide according to any one of claims 11, 15 or 17, comprising the steps of introducing an isolated nucleic acid according to any one of claims 1, 4, 5 or 8 or a vector according to claim 9 into a host
20 cell, culturing said host cell under conditions suitable for expression of said polypeptide, and recovering said polypeptide.

23. An antibody that binds to a polypeptide of any one of claims 11, 15
25 or 17.

24. A composition comprising the polypeptide of claim 11, 15 or 17 and a pharmaceutically acceptable carrier.

25. A composition comprising the antibody of claim 23 and a pharmaceutically acceptable carrier.
30

- 142 -

26. A method of decreasing inhibition of axonal growth of a CNS neuron, comprising the step of contacting the neuron with an effective amount of the polypeptide of claim 11, 15 or 17.

5 27. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the polypeptide of claim 11, 15 or 17.

10 28. A method of decreasing inhibition of axonal growth of a CNS neuron comprising the step of contacting the neuron with an effective amount of the antibody according to claim 23.

15 29. A method of treating a central nervous system disease, disorder or injury, comprising administering to a mammal an effective amount of the antibody according to claim 23.

30. A method for identifying a molecule that binds a polypeptide of claim 11, 15 or 17 comprising the steps of:

- 20 (a) providing a polypeptide of claim 11, 15 or 17;
 (b) contacting the polypeptide with the candidate molecule;
 and
 (c) detecting binding of the candidate molecule to the polypeptide.

25

1/6

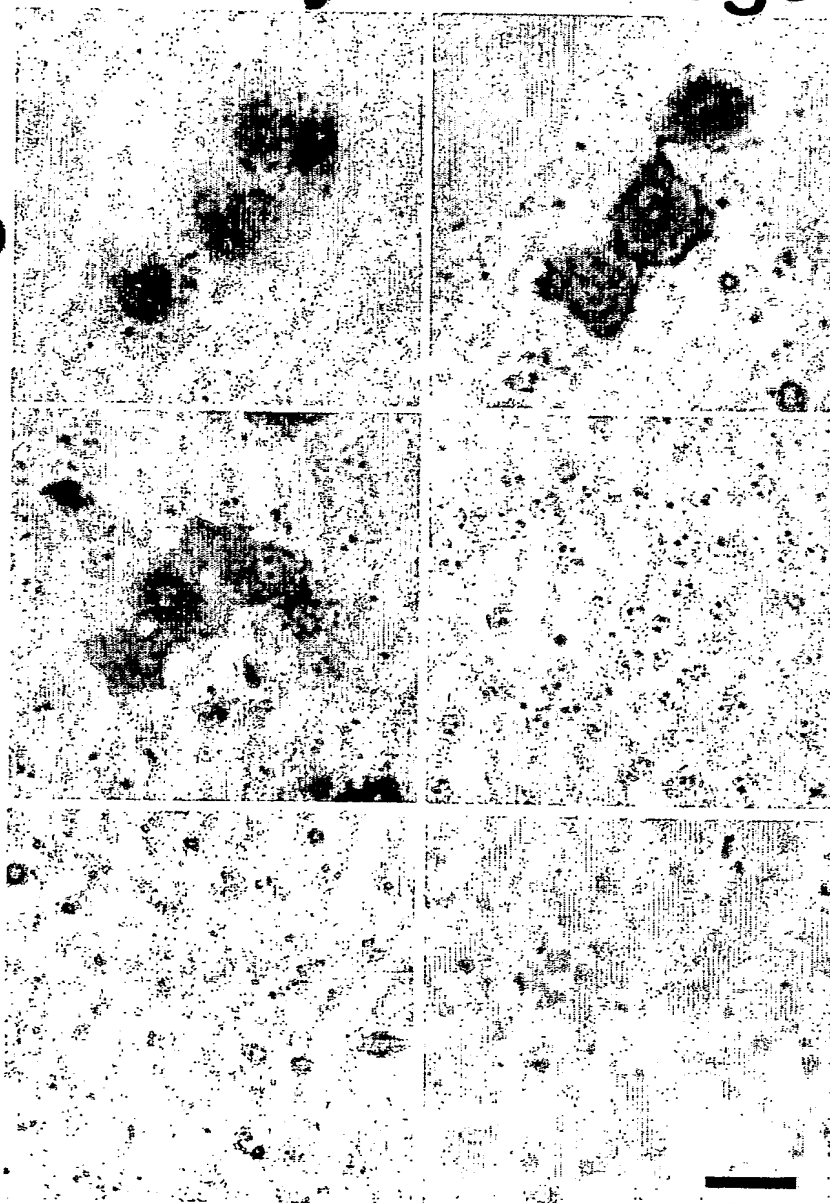
FIG. 1A

1	~~~~~MLPG	IRRLQAPAS	AC...LLML	LA..LPLAAP	SCPMLCTCYS	50
	MSWQSGTTVT	QSPVQAAQVS	GCCVELLLL	LAGEPLGG.	GCPRDCVCYP	
	~~~~~MKRAS	AGGSRLLA	WV	LWLQAWQVAA	PCPGACVCYN	
	~~~~~S	-----LL---	L-----	-CP--C-CY-		
	NOGO-R2					
	NOGO-R3					
	NOGO-R1					
	Consensus					
51	SP.PTVSCQA	NNFSSVPLSL	PPSTQRLFLQ	NNLIRTLRPG	TFGS..NLLT	100
	AP.MTVSCQA	HNFAAIEGI	PEDSERIFLQ	NNRITFLOQG	HFSP..AMVT	
	EPKVTTSQPQ	QGLQAVPVG	PAASQRIFLH	GNRISHVPAA	SFRACRNLT	
	-P--T-SC--	-----P---	P-----R-FL-	-N-I-----	-F-----	
	NOGO-R2					
	NOGO-R3					
	NOGO-R1					
	Consensus					
101	LWLFSSNNLST	IYPGTFRHLQ	ALAEELDLGDN	RHLRSLEPDT	FQGLERLQSL	150
	LWIYSNNITF	IAPNTFEGFV	HLEELDGLDN	RQLRTLAPET	FQGLVKLHAL	
	LWLHSNVLAR	IDAAAFGLA	LLEQLDLSN	AQLRSVDPAT	FHGLGRHLTL	
	LW--SN----	I-----F----	-LE-LDL-DN	--LR----P-T	F-GL--L--L	
	NOGO-R2					
	NOGO-R3					
	NOGO-R1					
	Consensus					
151	HLRYCQLSSL	PGNIFRGLVS	LQYLYQENS	LLHLQDDDLFA	DLANLSHLFL	200
	YLYKCGLSAL	PAGIFGGLHS	LQYLYQDNH	IEYLQDDIFV	DLVNLSHLFL	
	HLDRCGIQEL	GPGLFRGLAA	LQYLYQDNA	LQALPDDTFR	DLGNLTHLFL	
	-L--C-L--L	----F-GL--	LQYLYQ-N-	---L-DD-F-	DL-NL-HLFL	
	NOGO-R2					
	NOGO-R3					
	NOGO-R1					
	Consensus					
201	HGNRLRLITE	HVFRGLGSLD	RLLLHGNRLQ	GVHRAAFRGL	SRLTILYLFN	250
	HGNKLWSLQ	GIFRGLVNL	RLLLHENQLQ	VVHHKAFHDL	HRLTTLFLFN	
	HGNRISSVPE	RAFRGLHSLD	RLLLHQNVA	HVHPHAFRDL	GRIMTYLFA	
	HGN-----	--FRGL--LD	RLLLH-N---	-VH--AF--L	-RL--L-LF-	
	NOGO-R2					
	NOGO-R3					
	NOGO-R1					
	Consensus					

SUBSTITUTE SHEET (RULE 26)

451	ACQAPPDSRG	PALSAGLPSP	LLCLLLLVPH	HL	491
NOGO-R2	PIRAPSGVQQ	ASSGTALGAP	LLAWILGLAV	TLR	
NOGO-R3	QAGSGGGGTG	DSEGSGLPS	LTCSLTPLGL	ALVLWTVLGP	C
NOGO-R1	-----	-----	-----	-----	
Consensus	-----	-----	-----	-----	

3/6

anti-myc AP-Nogo**FIG. 2****Control NgR 3 NgR 1**

SUBSTITUTE SHEET (RULE 26)

	1	50
Human NOGO-R1	~~~~~	GGSRLLAWVL
Murine NOGO-R1	~~~~~	~~~~~MKRASA
Murine NOGO-R3	~~~~~	~~~~~MKRASS
Human NOGO-R3	~~~~~	MSWQSGTTVT
Human NOGO-R2	~~~~~	~~~~~QSPVQAAQVS
Consensus	~~~~~	~~~~~MLPGLRRLIQ
	LRR NT	LRR 1
Human NOGO-R1	PKVTTSCPQQ	AASQRIFLHG
Murine NOGO-R1	PKVTTSCPQQ	ASSQRIFLHG
Murine NOGO-R3	P.MTVSCQAH	EDSERIFLQN
Human NOGO-R3	~~~~~	VDSEKVFELQN
Human NOGO-R2	P.PTVSCQAN	PSTQRLFLQN
Consensus	P--T-SC--	----R-FL--
	LRR 2	LRR 3
Human NOGO-R1	WLHSNVIARI	LEQLDLSDNA
Murine NOGO-R1	WLHSNALARE	LEQLDLSDNA
Murine NOGO-R3	WIYSNNITFI	LEELDLGDNR
Human NOGO-R3	WIYSNNITYI	LEELDLGDNR
Human NOGO-R2	WLFNNLSTI	LEELDLGDNR
Consensus	W--SN--	LE-LDL-DN-
	LRR 4	LRR 5
Human NOGO-R1	LDRCGLOELG	QYLYIQDNAL
Murine NOGO-R1	LDRCGLRELG	QYLYIQDNAL
Murine NOGO-R3	LYKCGLSALP	QYLYIQDNHI
Human NOGO-R3	LYKCGLSALP	QYLYIQDNHI
Human NOGO-R2	LYRCQLSSLP	QYLYIQENS
Consensus	L--C-L-L-	QYLYIQ-N--
	LRR 6	LRR 7
Human NOGO-R1	WLRGRLHLLH	WLRGRLHLLH
Murine NOGO-R1	WLRGRLHLLH	WLRGRLHLLH
Murine NOGO-R3	WLRGRLHLLH	WLRGRLHLLH
Human NOGO-R3	WLRGRLHLLH	WLRGRLHLLH
Human NOGO-R2	WLRGRLHLLH	WLRGRLHLLH
Consensus	WLRGRLHLLH	WLRGRLHLLH

FIG. 3, cont.

SUBSTITUTE SHEET (RULE 26)

6/6

FIG. 3, cont.

Human	NOGO-R1	401	LPGSAEPPLT	AVRPEGSEPP	GF...PTSGP	RRRPGCSRKN	RTRSHCRLGQ	450
Murine	NOGO-R1		LPSSAEPPLT	ALRPGGSEPP	GL...PTTGP	RRRPGCSRKN	RTRSHCRLGQ	
Murine	NOGO-R3		.SGKELTELQ	DYAPDYQHKF	SFDIMPTARP	KRKGKCARRT	PIRAPSGVQQ	
Human	NOGO-R3		.AGKQAPPELP	DYAPDYQHKF	SFDIMPTARP	KRKGKCARRT	PIRAPSGVQQ	
Human	NOGO-R2		.QGGDAPTED	DYWGGY....GGED	QRGEQMCPGA	ACQAPPD...	
Consensus			-----	-----	-----	-----	-----	
						-R		

Human	NOGO-R1	451	AGSGGGGTGD	SE	Putative GPI Signals	490
Murine	NOGO-R1		AGSGASGTGD	SE	CGSGALPSL	TCSLTPLGLA
Murine	NOGO-R3		AE	CGSGALPAL	ACSLAPLGLA
Human	NOGO-R3		ASSGTALG	APLLAWITGL
Human	NOGO-R3		ASSASSLG	ASLLAWITGL
Human	NOGO-R2		SRGPALSA	GLPS
Consensus			-----	---	-----	-----

SEQUENCE LISTING

<110> BIOGEN, INC.
YALE UNIVERSITY
STRITTMATTER, STEPHEN M.
CATE, RICHARD L.
SAH, DINAH W.Y.

<120> NOGO RECEPTOR HOMOLOGS

<130> A116PCT

<140>

<141>

<150> 60/238,361

<151> 2000-10-06

<160> 16

<170> PatentIn Ver. 2.1

<210> 17

<211> 1260

<212> DNA

<213> Homo sapiens

<400> 1

```

atgctgcccc ggctcaggcg cctgctgcaa gctcccgctt cggcctgcct cctgctgatg 60
ctcctggccc tgccctggc ggcccccagc tgcccatgct tctgcacctg ctactcatcc 120
ccgcccaccg tgagctgcca ggccaacaac ttctcctctg tgccgctgtc cctgccaccc 180
agcactcagc gactcttcct gcagaacaac ctcatccgca cgtgcgggcc aggcaccttt 240
gggtccaacc tgctcaccct gtggctcttc tccaacaacc tctccacctt ctaccggggc 300
actttccgcc acttgcaagc cctggaggag ctggacctcg gtgacaaccg gcacctgcgc 360
tcgctggagc ccgacacctt ccagggcctg gaggcgctgc agtcgctgca tttgtaccgc 420
tgccagctca gcagcctgcc cggcaacatc ttccgaggcc tggtcagcct gcagtacctc 480
tacctccagg agaacagcct gctccacctt caggatgact tggtcgcgga cctggccaac 540
ctgagccacc tcttcctcca cgggaaccgc ctgcggtgc tcacagagca cgtgtttcgc 600
ggcctgggca gcctggaccg gctgctgctg cagggaacc ggctgcaggc cgtgcaccgc 660
gcggccttcc gcggcctcag ccgcctcacc atcctctacc tgttcaacaa cagcctggcc 720
tcgctgcccg gcgaggcgct cgccgacctg ccctcgctcg agttcctgcy gctcaacgct 780
aaccctggg cgtgcgactg ccgcgcgcgg ccgctctggg cctgggtcca gcgcgcgcgc 840
gtgtccagct ccgacgtgac ctgcgccacc ccccgaggc gccaggggcc agacctgcgc 900
gcgtcccgcg aggcgactt ccaggcgtgt ccgcccgcgg caccacgcgc gccgggcagc 960
cgcgcccgcg gcaacagctc ctccaaccac ctgtacgggg tggccgaggc cggggcgccc 1020
ccagccgatc cctccaccct ctaccgagat ctgcctgccg aagactcgcg ggggcgccag 1080
ggcggggacg cgcctactga ggacgactac tgggggggct acgggggtga ggaccagcga 1140
ggggagcaga tgtgccccgg cgctgcctgc caggcgcccc cggactcccg aggcctgcg 1200
ctctcgcccg ggctccccag ccctctgctt tgccctctgc tctgggtgcc ccaccacctc 1260

```

<210> 2

<211> 420

<212> PRT

<213> Homo sapiens

<400> 2

```

Met Leu Pro Gly Leu Arg Arg Leu Leu Gln Ala Pro Ala Ser Ala Cys
 1             5             10             15

```

```

Leu Leu Leu Met Leu Leu Ala Leu Pro Leu Ala Ala Pro Ser Cys Pro
 20             25             30

```

Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys Gln Ala
 35 40 45
 Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr Gln Arg
 50 55 60
 Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly Thr Phe
 65 70 75 80
 Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu Ser Thr
 85 90 95
 Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu Leu Asp
 100 105 110
 Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr Phe Gln
 115 120 125
 Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln Leu Ser
 130 135 140
 Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln Tyr Leu
 145 150 155 160
 Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu Phe Ala
 165 170 175
 Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg Leu Arg
 180 185 190
 Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp Arg Leu
 195 200 205
 Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala Phe Arg
 210 215 220
 Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser Leu Ala
 225 230 235 240
 Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu Phe Leu
 245 250 255
 Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg Pro Leu
 260 265 270
 Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val Thr Cys
 275 280 285
 Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu Arg Glu
 290 295 300
 Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro Gly Ser
 305 310 315 320
 Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val Ala Glu
 325 330 335
 Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp Leu Pro
 340 345 350
 Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr Glu Asp
 355 360 365
 Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu Gln Met

370 375 380

Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly Pro Ala
 385 390 395 400

Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu Leu Val
 405 410 415

Pro His His Leu
 420

<210> 3
 <211> 1383
 <212> DNA
 <213> Mus sp.

<400> 3

atgtcttggc	agtctggaac	cacagtgaca	caatctcccg	tgcaggctgc	tcagggtctca	60
gggtgctgtg	tggaattgct	gctgttgctg	ctcgtctggag	agctacctct	gggtgggtgt	120
tgtcctcgag	actgtgtgtg	ctaccctgcg	cccatgactg	tcagctgcca	ggcacacaac	180
tttgctgcca	tcccggaggg	catcccagag	gacagtggag	gcatcttcct	gcagaacaat	240
cgcacacact	tcctccagca	gggccacttc	agccccgcca	tggtcaccct	ctggatctac	300
tccaacaaca	tcactttcat	tgctcccaac	accttcgagg	gctttgtgca	tctggaggag	360
ctagaccttg	gagacaaccg	acagctgcca	acgctggcac	ccgagacctt	ccaaggcctg	420
gtgaagcttc	acgccctcta	cctctataag	tgtggactga	gcgccctgcc	cgcaggcatc	480
tttgggtggc	tgcacagcct	gcagtatctc	tacttgagg	acaaccatat	cgagtacctc	540
caagatgaca	tctttgtgga	cctgggtcaat	ctcagtcact	tgtttctcca	tggttaacaag	600
ctatggagcc	tgggccaagg	catcttccgg	ggcctgggtga	acctggaccg	gttgctgctg	660
catgagaacc	agctacagtg	ggttcaccac	aaggctttcc	atgacctcca	caggctaacc	720
accctctttc	tcttcaacaa	cagcctcact	gagctgcagg	gtgactgtct	ggcccccttg	780
gtggccttgg	agttccttcg	cctcaatggg	aatgcttggg	actgtggctg	ccgggcacgt	840
tcctctgtgg	aatggctgcg	aagggtccgt	ggctctagct	ctgctgtccc	ctgcgcgacc	900
cccgagctgc	ggcaaggcca	ggatctgaag	ctgctgaggg	tggaggactt	ccggaactgc	960
acaggaccag	tgtctcctca	ccagatcaag	tctcacacgc	ttaccacctc	tgacagggct	1020
gcccgaagg	agcaccatcc	gtcccatggg	gcctccaggg	acaaaggcca	cccacatggc	1080
catccgcctg	gctccaggtc	aggttacaag	aaggcaggca	agaactgcac	cagccacagg	1140
aaccggaacc	agatctctaa	ggtgagctct	gggaaagagc	ttaccgaact	gcaggactat	1200
gcccccgact	atcagcacaa	gttcagcttt	gacatcatgc	ccaccgcacg	acccaagagg	1260
aagggcaagt	gtgctcgag	gacccccatc	cgtgccccca	gtgggggtgca	gcaggcatcc	1320
tcaggcacgg	cccttggggc	cccactcctg	gcctggatac	tgggggtggc	agtcactctc	1380
cgc						1383

<210> 4
 <211> 461
 <212> PRT
 <213> Mus sp.

<400> 4

Met	Ser	Trp	Gln	Ser	Gly	Thr	Thr	Val	Thr	Gln	Ser	Pro	Val	Gln	Ala
1				5					10					15	
Ala	Gln	Val	Ser	Gly	Cys	Cys	Val	Glu	Leu	Leu	Leu	Leu	Leu	Leu	Ala
			20					25					30		
Gly	Glu	Leu	Pro	Leu	Gly	Gly	Gly	Cys	Pro	Arg	Asp	Cys	Val	Cys	Tyr
		35					40					45			
Pro	Ala	Pro	Met	Thr	Val	Ser	Cys	Gln	Ala	His	Asn	Phe	Ala	Ala	Ile
	50					55					60				
Pro	Glu	Gly	Ile	Pro	Glu	Asp	Ser	Glu	Arg	Ile	Phe	Leu	Gln	Asn	Asn
65					70					75					80

Arg Ile Thr Phe Leu Gln Gln Gly His Phe Ser Pro Ala Met Val Thr
 85 90 95
 Leu Trp Ile Tyr Ser Asn Asn Ile Thr Phe Ile Ala Pro Asn Thr Phe
 100 105 110
 Glu Gly Phe Val His Leu Glu Glu Leu Asp Leu Gly Asp Asn Arg Gln
 115 120 125
 Leu Arg Thr Leu Ala Pro Glu Thr Phe Gln Gly Leu Val Lys Leu His
 130 135 140
 Ala Leu Tyr Leu Tyr Lys Cys Gly Leu Ser Ala Leu Pro Ala Gly Ile
 145 150 155 160
 Phe Gly Gly Leu His Ser Leu Gln Tyr Leu Tyr Leu Gln Asp Asn His
 165 170 175
 Ile Glu Tyr Leu Gln Asp Asp Ile Phe Val Asp Leu Val Asn Leu Ser
 180 185 190
 His Leu Phe Leu His Gly Asn Lys Leu Trp Ser Leu Gly Gln Gly Ile
 195 200 205
 Phe Arg Gly Leu Val Asn Leu Asp Arg Leu Leu Leu His Glu Asn Gln
 210 215 220
 Leu Gln Trp Val His His Lys Ala Phe His Asp Leu His Arg Leu Thr
 225 230 235 240
 Thr Leu Phe Leu Phe Asn Asn Ser Leu Thr Glu Leu Gln Gly Asp Cys
 245 250 255
 Leu Ala Pro Leu Val Ala Leu Glu Phe Leu Arg Leu Asn Gly Asn Ala
 260 265 270
 Trp Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Arg Arg
 275 280 285
 Phe Arg Gly Ser Ser Ser Ala Val Pro Cys Ala Thr Pro Glu Leu Arg
 290 295 300
 Gln Gly Gln Asp Leu Lys Leu Leu Arg Val Glu Asp Phe Arg Asn Cys
 305 310 315 320
 Thr Gly Pro Val Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr
 325 330 335
 Ser Asp Arg Ala Ala Arg Lys Glu His His Pro Ser His Gly Ala Ser
 340 345 350
 Arg Asp Lys Gly His Pro His Gly His Pro Pro Gly Ser Arg Ser Gly
 355 360 365
 Tyr Lys Lys Ala Gly Lys Asn Cys Thr Ser His Arg Asn Arg Asn Gln
 370 375 380
 Ile Ser Lys Val Ser Ser Gly Lys Glu Leu Thr Glu Leu Gln Asp Tyr
 385 390 395 400
 Ala Pro Asp Tyr Gln His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala
 405 410 415
 Arg Pro Lys Arg Lys Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala

			420						425										430
Pro	Ser	Gly	Val	Gln	Gln	Ala	Ser	Ser	Gly	Thr	Ala	Leu	Gly	Ala	Pro				
		435					440					445							
Leu	Leu	Ala	Trp	Ile	Leu	Gly	Leu	Ala	Val	Thr	Leu	Arg							
		450				455					460								
<p><210> 5</p> <p><211> 473</p> <p><212> PRT</p> <p><213> Homo sapiens</p>																			
<p><400> 5</p>																			
Met	Lys	Arg	Ala	Ser	Ala	Gly	Gly	Ser	Arg	Leu	Leu	Ala	Trp	Val	Leu				
1				5					10					15					
Trp	Leu	Gln	Ala	Trp	Gln	Val	Ala	Ala	Pro	Cys	Pro	Gly	Ala	Cys	Val				
			20					25					30						
Cys	Tyr	Asn	Glu	Pro	Lys	Val	Thr	Thr	Ser	Cys	Pro	Gln	Gln	Gly	Leu				
		35					40					45							
Gln	Ala	Val	Pro	Val	Gly	Ile	Pro	Ala	Ala	Ser	Gln	Arg	Ile	Phe	Leu				
	50					55					60								
His	Gly	Asn	Arg	Ile	Ser	His	Val	Pro	Ala	Ala	Ser	Phe	Arg	Ala	Cys				
	65				70					75					80				
Arg	Asn	Leu	Thr	Ile	Leu	Trp	Leu	His	Ser	Asn	Val	Leu	Ala	Arg	Ile				
				85					90					95					
Asp	Ala	Ala	Ala	Phe	Thr	Gly	Leu	Ala	Leu	Leu	Glu	Gln	Leu	Asp	Leu				
			100					105					110						
Ser	Asp	Asn	Ala	Gln	Leu	Arg	Ser	Val	Asp	Pro	Ala	Thr	Phe	His	Gly				
		115					120					125							
Leu	Gly	Arg	Leu	His	Thr	Leu	His	Leu	Asp	Arg	Cys	Gly	Leu	Gln	Glu				
	130					135					140								
Leu	Gly	Pro	Gly	Leu	Phe	Arg	Gly	Leu	Ala	Ala	Leu	Gln	Tyr	Leu	Tyr				
	145				150					155					160				
Leu	Gln	Asp	Asn	Ala	Leu	Gln	Ala	Leu	Pro	Asp	Asp	Thr	Phe	Arg	Asp				
				165					170					175					
Leu	Gly	Asn	Leu	Thr	His	Leu	Phe	Leu	His	Gly	Asn	Arg	Ile	Ser	Ser				
			180					185					190						
Val	Pro	Glu	Arg	Ala	Phe	Arg	Gly	Leu	His	Ser	Leu	Asp	Arg	Leu	Leu				
		195					200					205							
Leu	His	Gln	Asn	Arg	Val	Ala	His	Val	His	Pro	His	Ala	Phe	Arg	Asp				
	210					215					220								
Leu	Gly	Arg	Leu	Met	Thr	Leu	Tyr	Leu	Phe	Ala	Asn	Asn	Leu	Ser	Ala				
	225				230					235					240				
Leu	Pro	Thr	Glu	Ala	Le														

260						265						270					
Ala	Trp	Leu	Gln	Lys	Phe	Arg	Gly	Ser	Ser	Ser	Glu	Val	Pro	Cys	Ser		
		275					280					285					
Leu	Pro	Gln	Arg	Leu	Ala	Gly	Arg	Asp	Leu	Lys	Arg	Leu	Ala	Ala	Asn		
	290					295					300						
Asp	Leu	Gln	Gly	Cys	Ala	Val	Ala	Thr	Gly	Pro	Tyr	His	Pro	Ile	Trp		
305					310					315					320		
Thr	Gly	Arg	Ala	Thr	Asp	Glu	Glu	Pro	Leu	Gly	Leu	Pro	Lys	Cys	Cys		
				325					330					335			
Gln	Pro	Asp	Ala	Ala	Asp	Lys	Ala	Ser	Val	Leu	Glu	Pro	Gly	Arg	Pro		
			340					345					350				
Ala	Ser	Ala	Gly	Asn	Ala	Leu	Lys	Gly	Arg	Val	Pro	Pro	Gly	Asp	Ser		
		355					360						365				
Pro	Pro	Gly	Asn	Gly	Ser	Gly	Pro	Arg	His	Ile	Asn	Asp	Ser	Pro	Phe		
	370					375					380						
Gly	Thr	Leu	Pro	Gly	Ser	Ala	Glu	Pro	Pro	Leu	Thr	Ala	Val	Arg	Pro		
385					390					395					400		
Glu	Gly	Ser	Glu	Pro	Pro	Gly	Phe	Pro	Thr	Ser	Gly	Pro	Arg	Arg	Arg		
				405					410					415			
Pro	Gly	Cys	Ser	Arg	Lys	Asn	Arg	Thr	Arg	Ser	His	Cys	Arg	Leu	Gly		
			420					425					430				
Gln	Ala	Gly	Ser	Gly	Gly	Gly	Gly	Thr	Gly	Asp	Ser	Glu	Gly	Ser	Gly		
		435					440					445					
Ala	Leu	Pro	Ser	Leu	Thr	Cys	Ser	Leu	Thr	Pro	Leu	Gly	Leu	Ala	Leu		
	450					455					460						
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys									
465					470												

<210> 6

<211> 440

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus sequence

<220>

<221> MOD_RES

<222> (3)..(4)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (6)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (9)..(10)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (12)..(13)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (15)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (18)..(25)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (27)..(29)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (31)..(34)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (36)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (39)..(40)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (42)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (44)..(50)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (52)..(59)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (62)..(63)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (66)..(69)

<223> Variable amino acid

<220>

<221> MOD_RES

<222> (71)..(74)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (76)..(80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (87)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)..(91)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (94)..(96)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (98)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (101)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(105)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (111)..(112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>

<221> MOD_RES
<222> (116)..(117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (119)..(122)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (124)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (127)..(128)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (136)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (138)..(141)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (143)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (146)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (148)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (151)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (154)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162)..(170)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (175)..(176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (184)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (186)..(189)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (192)..(193)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (196)..(197)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (199)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (202)..(203)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (205)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (208)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (210)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (212)..(213)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (215)..(218)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (221)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (223)..(224)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (226)..(227)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (232)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (234)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (236)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (238)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (243)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (246)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (248)..(251)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (253)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (257)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (259)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (261)..(262)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (264)..(267)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (269)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (272)..(273)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (275)..(277)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (279)..(281)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (283)..(287)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (289)..(290)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (292)..(328)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (330)..(341)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (344)..(346)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (348)..(399)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (401)..(428)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (431)..(439)
 <223> Variable amino acid

<400> 6
 Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
 1 5 10 15

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Pro Xaa Xaa
 20 25 30
 Xaa Xaa Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa
 35 40 45
 Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser
 50 55 60
 Asn Xaa Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa
 65 70 75 80
 Leu Glu Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa
 85 90 95
 Pro Xaa Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa
 100 105 110
 Cys Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa
 115 120 125
 Leu Gln Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp
 130 135 140
 Asp Xaa Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly
 145 150 155 160
 Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa
 165 170 175
 Leu Asp Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa
 180 185 190
 Xaa Ala Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa
 195 200 205
 Asn Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa
 210 215 220
 Leu Xaa Xaa Leu Arg Leu Asn Xaa Asn Xaa Trp Xaa Cys Xaa Cys Arg
 225 230 235 240
 Ala Arg Xaa Leu Trp Xaa Trp Xaa Xaa Xaa Xaa Arg Xaa Ser Ser Ser
 245 250 255
 Xaa Val Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Gly Xaa Asp Leu Xaa
 260 265 270
 Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Pro
 275 280 285
 Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 290 295 300
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 305 310 315 320
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 325 330 335
 Xaa Xaa Xaa Xaa Xaa Pro Pro Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa
 340 345 350
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

355 360 365
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 370 375 380
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg
 385 390 395 400
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 405 410 415
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 420 425 430
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 435 440

<210> 7
 <211> 66
 <212> PRT
 <213> Homo sapiens

<400> 7
 Arg Ile Tyr Lys Gly Val Ile Gln Ala Ile Gln Lys Ser Asp Glu Gly
 1 5 10 15
 His Pro Phe Arg Ala Tyr Leu Glu Ser Glu Val Ala Ile Ser Glu Glu
 20 25 30
 Leu Val Gln Lys Tyr Ser Asn Ser Ala Leu Gly His Val Asn Cys Thr
 35 40 45
 Ile Lys Glu Leu Arg Arg Leu Phe Leu Val Asp Asp Leu Val Asp Ser
 50 55 60
 Leu Lys
 65

<210> 8
 <211> 390
 <212> PRT
 <213> Homo sapeins

<400> 8
 Cys Pro Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys
 1 5 10 15
 Gln Ala Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr
 20 25 30
 Gln Arg Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly
 35 40 45
 Thr Phe Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu
 50 55 60
 Ser Thr Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr
 85 90 95

Phe Gln Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln
 100 105 110
 Leu Ser Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu
 130 135 140
 Phe Ala Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg
 145 150 155 160
 Leu Arg Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp
 165 170 175
 Arg Leu Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala
 180 185 190
 Phe Arg Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser
 195 200 205
 Leu Ala Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg
 225 230 235 240
 Pro Leu Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val
 245 250 255
 Thr Cys Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu
 260 265 270
 Arg Glu Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro
 275 280 285
 Gly Ser Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val
 290 295 300
 Ala Glu Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp
 305 310 315 320
 Leu Pro Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr
 325 330 335
 Glu Asp Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu
 340 345 350
 Gln Met Cys Pro Gly Ala Ala Cys Gln Ala Pro Pro Asp Ser Arg Gly
 355 360 365
 Pro Ala Leu Ser Ala Gly Leu Pro Ser Pro Leu Leu Cys Leu Leu Leu
 370 375 380
 Leu Val Pro His His Leu
 385 390

<210> 9
 <211> 421
 <212> PRT
 <213> Mus sp.

<400> 9

Cys Pro Arg Asp Cys Val Cys Tyr Pro Ala Pro Met Thr Val Ser Cys
 1 5 10 15
 Gln Ala His Asn Phe Ala Ala Ile Pro Glu Gly Ile Pro Glu Asp Ser
 20 25 30
 Glu Arg Ile Phe Leu Gln Asn Asn Arg Ile Thr Phe Leu Gln Gln Gly
 35 40 45
 His Phe Ser Pro Ala Met Val Thr Leu Trp Ile Tyr Ser Asn Asn Ile
 50 55 60
 Thr Phe Ile Ala Pro Asn Thr Phe Glu Gly Phe Val His Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg Gln Leu Arg Thr Leu Ala Pro Glu Thr
 85 90 95
 Phe Gln Gly Leu Val Lys Leu His Ala Leu Tyr Leu Tyr Lys Cys Gly
 100 105 110
 Leu Ser Ala Leu Pro Ala Gly Ile Phe Gly Gly Leu His Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Asp Asn His Ile Glu Tyr Leu Gln Asp Asp Ile
 130 135 140
 Phe Val Asp Leu Val Asn Leu Ser His Leu Phe Leu His Gly Asn Lys
 145 150 155 160
 Leu Trp Ser Leu Gly Gln Gly Ile Phe Arg Gly Leu Val Asn Leu Asp
 165 170 175
 Arg Leu Leu Leu His Glu Asn Gln Leu Gln Trp Val His His Lys Ala
 180 185 190
 Phe His Asp Leu His Arg Leu Thr Thr Leu Phe Leu Phe Asn Asn Ser
 195 200 205
 Leu Thr Glu Leu Gln Gly Asp Cys Leu Ala Pro Leu Val Ala Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Gly Asn Ala Trp Asp Cys Gly Cys Arg Ala Arg
 225 230 235 240
 Ser Leu Trp Glu Trp Leu Arg Arg Phe Arg Gly Ser Ser Ser Ala Val
 245 250 255
 Pro Cys Ala Thr Pro Glu Leu Arg Gln Gly Gln Asp Leu Lys Leu Leu
 260 265 270
 Arg Val Glu Asp Phe Arg Asn Cys Thr Gly Pro Val Ser Pro His Gln
 275 280 285
 Ile Lys Ser His Thr Leu Thr Thr Ser Asp Arg Ala Ala Arg Lys Glu
 290 295 300
 His His Pro Ser His Gly Ala Ser Arg Asp Lys Gly His Pro His Gly
 305 310 315 320
 His Pro Pro Gly Ser Arg Ser Gly Tyr Lys Lys Ala Gly Lys Asn Cys
 325 330 335
 Thr Ser His Arg Asn Arg Asn Gln Ile Ser Lys Val Ser Ser Gly Lys

340				345				350							
Glu	Leu	Thr	Glu	Leu	Gln	Asp	Tyr	Ala	Pro	Asp	Tyr	Gln	His	Lys	Phe
355				360				365							
Ser	Phe	Asp	Ile	Met	Pro	Thr	Ala	Arg	Pro	Lys	Arg	Lys	Gly	Lys	Cys
370				375				380							
Ala	Arg	Arg	Thr	Pro	Ile	Arg	Ala	Pro	Ser	Gly	Val	Gln	Gln	Ala	Ser
385				390				395				400			
Ser	Gly	Thr	Ala	Leu	Gly	Ala	Pro	Leu	Leu	Ala	Trp	Ile	Leu	Gly	Leu
				405				410				415			
Ala	Val	Thr	Leu	Arg											
420															

<210> 10

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Consensus
sequence

<220>

<221> MOD RES

<222> (3) . . (4)

<223> Variable amino acid

<220>

<221> MOD RES

<222> (6)

<223> Variable amino acid

<220>

<221> MOD RES

<222> (9) . . (10)

<223> Variable amino acid

<220>

<221> MOD RES

$\langle 222 \rangle \quad (12) \bar{.} (13)$

<223> Variable amino acid

 $\langle 220 \rangle$

<221> MOD RES

 $\langle 222 \rangle \quad (15)^{-}$

<223> Variable amino acid

<400> 10

Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
1 5 10 15

Cys

<210> 11

<211> 50

<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Consensus
sequence

<220>
<221> MOD_RES
<222> (2)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (4)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (6)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (11)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (14)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (16)..(19)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (21)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (25)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (27)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (29)..(30)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (40)..(41)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (43)..(45)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (47)..(49)
 <223> Variable amino acid

<400> 11
 Asn Xaa Trp Xaa Cys Xaa Cys Arg Ala Arg Xaa Leu Trp Xaa Trp Xaa
 1 5 10 15
 Xaa Xaa Xaa Arg Xaa Ser Ser Ser Xaa Val Xaa Cys Xaa Xaa Pro Xaa
 20 25 30
 Xaa Xaa Xaa Gly Xaa Asp Leu Xaa Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa
 35 40 45
 Xaa Cys
 50

<210> 12
 <211> 196
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Consensus
 sequence

<220>
 <221> MOD_RES
 <222> (2)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (5)..(6)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (8)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (10)..(16)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (18)..(25)
 <223> Variable amino acid

<220>

<221> MOD_RES
<222> (28)..(29)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (37)..(40)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (42)..(46)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (49)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (53)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (56)..(57)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (60)..(62)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (64)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (67)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (70)..(71)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (73)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (75)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (77)..(78)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (80)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (82)..(83)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (85)..(88)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (90)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (93)..(94)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (102)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (104)..(107)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (109)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (112)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (114)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (117)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (120)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (128)..(136)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (141)..(142)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (150)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (152)..(155)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (158)..(159)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (162)..(163)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (165)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (168)..(169)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (171)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (174)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (176)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (178)..(179)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (181)..(184)
<223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (187)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (189)..(190)
 <223> Variable amino acid

<220>
 <221> MOD_RES
 <222> (192)..(193)
 <223> Variable amino acid

<400> 12

```

Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1           5           10           15
Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser Asn Xaa
          20           25           30
Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Leu Glu
          35           40           45
Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa Pro Xaa
          50           55           60
Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa Cys Xaa
 65           70           75           80
Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa Leu Gln
          85           90           95
Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp Asp Xaa
          100          105          110
Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly Asn Xaa
          115          120          125
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa Leu Asp
          130          135          140
Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa Xaa Ala
          145          150          155          160
Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa Asn Xaa
          165          170          175
Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa Leu Xaa
          180          185          190
Xaa Leu Arg Leu
          195

```

<210> 13
 <211> 1176
 <212> DNA
 <213> Homo sapiens

<400> 13
 gagggcatcc ccgtggacag cgagcgcgtc ttctctgcaga acaaccgcat cggcctcctc 60
 cagcccggcc acttcagccc cgccatgggc accctgtgga tctactcgaa caacatcacc 120

```

tacatccacc ccagcacctt cgagggttc gtgcacctgg aggagctgga cctcggcgac 180
aaccggcagc tgcggacgct ggcacccgag accttcagg gcctggtgaa gcttcacgcc 240
ctctacctct acaagtgtgg gctcagcgcc ttgccggccg gcgtctttgg cggcctgcac 300
agcctgcagt acctctacct gcaggacaac cacatcgagt acctccagga cgacatcttc 360
gtggacctgg tcaacctcag ccacctgttt ctccacggca acaagctgtg gactctgggc 420
ccgggcacct tccggggcct ggtgaacctg gacctcttt tgctgcacga gaaccagctg 480
cagtgggtcc accacaaggc attccacgac ctccgcaggc tgaccacctt ctctctcttc 540
aacaacagcc tctcggagct gcagggtgag tgcctggccc cgctgggggc cctggagttc 600
ctccgctca acggcaaccc ctgggactgt ggttgctcgc cgcgtcccct gtgggaatgg 660
ctgcagaggt tccggggctc cagctccgct gtcccctgtg tgcccctgg gctgcggcac 720
ggccaggacc tgaagctgct gagggccgag gacttccgga actgcacggg accagcgtcc 780
ccgcaccaga tcaagtcaca cacgtcacc accaccgaca gggccgcccg caaggaacac 840
cactcacccc acggcccccac caggagcaag ggccacccgc acggcccccg gcccggccac 900
aggaagccgg ggaagaactg caccaacccc aggaaccgca atcagatctc taaggcgggc 960
gccgggaaac agggcccccga gctgccagac tatgcccag actaccagca caagtccagt 1020
tttgacatca tgcctacggc ccggcccaag aggaagggca agtgtgccg caggaccccc 1080
atccgtgccc ccagcggggt gcagcaggcc tctcggcca gttccctggg ggcctccctc 1140
ctggcctgga cactggggct ggcggtcact ctccgc 1176

```

<210> 14

<211> 392

<212> PRT

<213> Homo sapiens

<400> 14

```

Glu Gly Ile Pro Val Asp Ser Glu Arg Val Phe Leu Gln Asn Asn Arg
 1             5             10             15

Ile Gly Leu Leu Gln Pro Gly His Phe Ser Pro Ala Met Val Thr Leu
          20             25             30

Trp Ile Tyr Ser Asn Asn Ile Thr Tyr Ile His Pro Ser Thr Phe Glu
          35             40             45

Gly Phe Val His Leu Glu Glu Leu Asp Leu Gly Asp Asn Arg Gln Leu
 50             55             60

Arg Thr Leu Ala Pro Glu Thr Phe Gln Gly Leu Val Lys Leu His Ala
 65             70             75             80

Leu Tyr Leu Tyr Lys Cys Gly Leu Ser Ala Leu Pro Ala Gly Val Phe
          85             90             95

Gly Gly Leu His Ser Leu Gln Tyr Leu Tyr Leu Gln Asp Asn His Ile
          100            105            110

Glu Tyr Leu Gln Asp Asp Ile Phe Val Asp Leu Val Asn Leu Ser His
          115            120            125

Leu Phe Leu His Gly Asn Lys Leu Trp Ser Leu Gly Pro Gly Thr Phe
          130            135            140

Arg Gly Leu Val Asn Leu Asp Arg Leu Leu Leu His Glu Asn Gln Leu
          145            150            155            160

Gln Trp Val His His Lys Ala Phe His Asp Leu Arg Arg Leu Thr Thr
          165            170            175

Leu Phe Leu Phe Asn Asn Ser Leu Ser Glu Leu Gln Gly Glu Cys Leu
          180            185            190

Ala Pro Leu Gly Ala Leu Glu Phe Leu Arg Leu Asn Gly Asn Pro Trp
          195            200            205

```


Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Gln Arg Phe
 210 215 220
 Arg Gly Ser Ser Ser Ala Val Pro Cys Val Ser Pro Gly Leu Arg His
 225 230 235 240
 Gly Gln Asp Leu Lys Leu Leu Arg Ala Glu Asp Phe Arg Asn Cys Thr
 245 250 255
 Gly Pro Ala Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr Thr
 260 265 270
 Asp Arg Ala Ala Arg Lys Glu His His Ser Pro His Gly Pro Thr Arg
 275 280 285
 Ser Lys Gly His Pro His Gly Pro Arg Pro Gly His Arg Lys Pro Gly
 290 295 300
 Lys Asn Cys Thr Asn Pro Arg Asn Arg Asn Gln Ile Ser Lys Ala Gly
 305 310 315 320
 Ala Gly Lys Gln Ala Pro Glu Leu Pro Asp Tyr Ala Pro Asp Tyr Gln
 325 330 335
 His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala Arg Pro Lys Arg Lys
 340 345 350
 Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala Pro Ser Gly Val Gln
 355 360 365
 Gln Ala Ser Ser Ala Ser Ser Leu Gly Ala Ser Leu Leu Ala Trp Thr
 370 375 380
 Leu Gly Leu Ala Val Thr Leu Arg
 385 390

<210> 15

<211> 143899

<212> DNA

<213> Homo sapiens

<220>

<221> modified_base

<222> (2044)..(2144)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (6609)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (6625)..(6724)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (14153)..(14252)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (19512)..(19611)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (22595)..(22694)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (27825)..(27924)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (34953)..(35052)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (40783)..(40882)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (49000)..(49099)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (62884)..(62983)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (75528)..(75627)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (87944)..(88043)
 <223> a, t, c, g, other or unknown

 <220>
 <221> modified_base
 <222> (111030)..(111129)
 <223> a, t, c, g, other or unknown

 <400> 15
 aagcacatac aggtgacatt acagaactga cagttatgcc aggcactgta cttagcccct 60
 ataccatcct caaacagctg tatgatgtag attgggtatt aaccccatta ataacaaaag 120
 tacagggaaac aaagtgactt tccaaaggctc atgccattca aaggagggtg aatcttaggt 180
 tggacgcagg ctgtctgact ctggagctcg aggtgttaat gctgcctcct ccatgggaac 240
 agcccaagtg aaaaacagct gatccactct tcatcttactt ggcattctgtg ctaagctggt 300
 ccctgagcca agctctgagc aacagaaaca gaagctctgc attaggagct tgtgagcatg 360
 tcaatgccgg gtaaaggagt gctggaaacc gctgggatgg ccgccgagca ctaggccggtt 420
 gaaggtgggc tctgtgtgac tggttcctct acactctggc ctggtgcct gcaggaagaa 480
 gatcaagctg agtgggctgg ccctggacca caaggtgaca ggtgacctct tctacacca 540
 tgtgaccacc atgggccaga ggctcagcca gaaggccccc agcctggagg acggttcgga 600
 tgccttcattg tcaccccagg atgttcgggg cacctcagaa aaccttctct agagttagtg 660
 tctggtcaag gtgccggcct tgggggatag tgatgggtggg tcctcatatt cagttagcac 720
 tcatggttga gtatttatc gcacccctct tcagtcctta caacaccca tgatgtaggt 780
 ggggcatgct cctcatttac agatgggcac atcaaagctc agctaacgct ggggaagttca 840
 gattcagggt taccctgctg gattcctggg attggggagg gaggagcttc caaatgggg 900

acaaggtctc	tgggcctgtc	gggtagctgg	tttcctcagg	gccccttgca	acctctgagc	960
ttattgcatc	aggtgcagcc	aggcccgtag	gcctcctggc	aggggtcctc	cacacctggc	1020
tgtcttttgc	cccctgctgg	tcacaggagg	agctgcagca	cctgcctggg	ctgcttctca	1080
ggagggtaca	tgaagatccc	aggaccgcca	gctccatgat	aagtggagg	agctccttgg	1140
agtcaggagc	gggagttgag	gagtttgagt	cctgctctcc	agttataggc	tatgtgactt	1200
gtgtagatca	cctaacccttg	ctcttgattt	ccttacctct	taaactagca	ctaaaagcac	1260
cccacaaact	gtaaagttag	ttgtgatgat	tgaatgacac	catgggtgtg	gaagctcttt	1320
gtaaagtgca	aaacggtgtg	cagtttgagg	gtggttaccc	ccagtgccga	ttctcagagg	1380
gcaacatggc	taagggcacg	agctggagtt	aggctgacct	gctgcttcca	gccctgtgag	1440
cttgagcaag	tcattttaact	tcctgagctg	cagtttcctc	atcagtataa	tgtgataagg	1500
ataggggtgt	tgtaaagattt	tattaaatgg	ggtaataaat	gtcaagtatg	tagcccatag	1560
tgagtgtctc	agagtttttt	tcttttgttt	ctttccccc	cgccccgaga	tggagcctta	1620
ctctgttgcc	caggtgtggg	tgcagtggca	tgatcttggc	tcactgcaac	ctccgcctcc	1680
cgggttcaag	caattctcct	gcctcagcct	cccaaatagc	tgggactaca	ggcgtgcacc	1740
accatgctcg	gctaattttt	gtatctttag	tagagacggg	gtttcaccat	gttggccagg	1800
ctggtctcga	actcctgacc	tcagtatgct	cctgcctcag	ccccgaaaag	ttttgggatt	1860
acaagtgtga	gcccccggtg	cctgccaggt	tttttttttt	tttttttttt	tgtaaaacac	1920
ccacagggtg	ttgctgttgc	ctgggtgtga	gtgctgtagt	gcaatcatag	ttcactgcag	1980
ccttgacctc	ctgggctcaa	gtgatectcc	tgcctcagcc	tcctgagtag	ctgggaatac	2040
aggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnntttttgt	ttttaagttag	2160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	ttggtctaaa	ctcccaacct	2220
agacagggtt	ttcccaatgt	tggccaaggg	gtactgggat	tacaggcgtg	agacaccgtg	2280
acccacctca	gcctcccaaa	gtactgggat	tcagggtatt	ctctacagcg	atgcaagaac	2340
aggcttattt	tcttgataaa	ttaccagctc	gacccagggt	ggagagctta	agatgtcaag	2400
agcctaatac	atccaggctc	agcatcagtg	acctagggac	tcctctctct	ttctcttttt	2460
gtctgaatgc	cgcttccaca	cacctttggg	gaccagtagt	tgggtgatgt	ttcagagata	2520
cagtagaaga	tgttatcttc	tcctttctct	tcactgagcc	caagttccct	cgctctctca	2580
gtttttcagt	caagatatat	ttcagtggct	ataacacata	ttttctaagg	tgggtgtgag	2640
ggacttttatt	tccttgtttc	tagaagaggg	aggacaggat	agagcaatgg	tgagaccact	2700
attaagggag	ctgggtaccg	gtggtgcata	acacttatca	aatcagcaca	cagtgaggagt	2760
caaaaagcga	aaagttgacc	tgcgaggggt	ttggagagtg	caagttatag	aaaaccagcg	2820
ggaaggaatg	tcctcatca	gttacaatat	atccagcagc	tttgaggagg	tgaggcaggt	2880
cctggccggg	cgcggtgggt	catgcctata	ctgaccaacg	tgggtgaaacc	ccacctctac	2940
ggatcacgag	gtcaggagtt	caagaccagc	tgtgtgcctg	taactctgagc	tactcaggag	3000
taaaaataca	aaattagctg	ggcggtgggt	aggtggagtt	tgacgtgagc	cgagatcgca	3060
gctgaggcac	gagaatcact	tgaaccgggg	agactccatc	tcaaacgaaa	aaaaaaaaag	3120
ccactgcact	ccagcctggg	caacagagcg	agtaagaaga	tctattatat	tatccatctc	3180
aaagaaaacc	cagctctaac	tggcttaaac	ccaaggttgg	ttgatccagg	gcttaacgat	3240
aggcagcagc	aagcccagag	gtaggggact	ggcacctctg	tctgcagggc	cagcttcatc	3300
gtcatcaaag	acccagggtc	tttctgtctc	agttggctgc	tgggcccagca	gacgctgcct	3360
ctaagccaga	ttgttcttgt	cttgattaca	tggcccttcc	ccagtcattg	aatgaaagag	3420
gcctccctgt	tcattcttcag	aagtagaaag	acccacctga	agccaatgac	tgtcaccagg	3480
tttcttttct	gtctgggatt	gcttaggtcc	gggttctctga	gccagtcttg	ggcaaggagt	3540
aaggtaatat	acactgattg	tcttaagtca	cagggcagct	ctgcagatga	gatcaaaact	3600
gtgatactgt	catgattgtc	ttgggtcat	aggaaagaga	cattttggga	gatacaaaac	3660
caagctacat	tattctgaac	agtgggaagt	gtcacagtgt	cttgggttga	catcaaatgt	3720
acaatgtcta	tcccatatcc	ctaggtccag	tacctattca	tctgggtttt	gatttttttt	3780
agaaaaagaa	agactgtcca	tccatttate	caccagact	ggagtgcagt	ggtttgatca	3840
aaattttatt	ttaagacatt	ctcactctgt	caagtgacce	tcccatgctc	aagtgatcct	3900
tggctcatgg	cagcctcaac	ctcccaggct	aaaggtgcat	gccaccagcg	tcagttaatt	3960
cctacctcag	cctcccaagt	agctagaact	catgatgccc	atgctagtct	ggaattcctg	4020
tttgcatttt	ttgtagagat	gggttttctg	caaatgctg	ggattgtagg	catgagccac	4080
aactcaagca	atatgcctgc	ctttgcctcc	gaaacaacta	ctgggtgctg	agcacggggc	4140
tgctcctggc	tcattctgtt	aataatttat	agatctctgc	caagtctgtg	gtttcttaaa	4200
caggggctgg	agatctagca	gggaccaggc	tctttctgtt	ttttcttttt	ctttctttct	4260
ggttttgctc	ataattcccc	ttttcttttc	ctcactctgt	taccagggtc	ggagtgcagt	4320
ttctttcttt	tttttttttt	gagacagagt	ctcctgggtt	caagcgattc	tcctgcctca	4380
ggtgcatct	cagctcactg	caacctctgc	tgccaccatg	cccggcta	ttttgtgttt	4440
gcctcccgag	tagctgggac	tacaggcgcc	caggctgggt	ttgaactcct	gaccttgtga	4500
ttagtagaga	ctgggtttca	ccatattggc	gattacaggc	gtgagccacg	gcgcccagcc	4560
tcgccccgct	tcggcctccc	acagtgcctg	ctccatactc	tactcctcca	cctggcccac	4620
agtttccctt	ttcaatgagg	cctccctgac	caagacctcc	taacatacta	cgtaagttat	4680
tcagctctac	tttttcttcc	ccatagcact	tcttctctca	ctagaatgta	aactccatga	4740
ttatttacta	ggcttactgt	gtattgtctg				

gaatagaaat	ttttgccttt	ttatttagtg	tggtgtctgc	agcccctggc	ttagtccctg	4800
gcatacaaca	gtcactccac	ccacagttgc	tgaataagtg	actaaaggctc	cctgccctca	4860
tattgtttatg	agggagtggtg	catgtttgtta	gagaaaaatc	tgaggcaca	taaaattttta	4920
tagagtttaa	gttttctttt	ttaagcaatc	cacgaattgg	ggtagtttca	gaggtagttt	4980
ttcagtcagt	acgtattttca	atggcttcac	tgagcccaag	ttctttcacc	tctctaggac	5040
tttattttcct	tattttctaga	acggggataa	cacatagttc	ataaggcagt	tatgagagta	5100
agggagctgg	tatgggggtga	tgcataagga	caggatagag	cagtgggtgag	accgctcaga	5160
tgacaaaagcg	tcagagacca	gtattttacga	cggaaatgtg	gaagcatgat	aaagaaatta	5220
tttggggcgg	gcacaatgac	tcacaactaa	taaaactttg	ggaggccaag	gtgggaggat	5280
cacttgactt	gcagaaggctc	aaggctgcag	tgagctgtga	ttttgccact	gcactccagc	5340
ctgggtcaaca	gagtgagacc	ctggctcgaa	acgttatttg	attggttaca	gttatacagt	5400
tgcccttattt	ggtctattcc	atgtgaaagt	tcctagttct	ataattttta	gtttgttggc	5460
tgttttctgat	tgggttaagct	taagttttgt	tttcccttaa	tacagttaag	tgccccataa	5520
tgacatttttg	gtcaaggaca	gaccacatat	acagtggtgg	tcccataaga	ttataatgga	5580
gctgaaacat	tcctattgtc	tatggcgtag	tggtcctgat	gttgtagcgc	aatgcattag	5640
ttatatgttt	gtggcaatgc	tggtgtaaac	acacctactg	cactgccagt	gatataaaag	5700
aatagcacat	acagttatat	atagtacata	atatctgata	atgataatac	ataactatat	5760
tactggttta	tatatattact	atattattta	tctttatttt	atttttgaga	cagagtctca	5820
ttctgtcacc	caggctggag	tgcagtggcg	cgatcttggc	tcaccgcaac	ctccgcttcc	5880
tgggttcaag	tgattctcct	gcctcagctc	cctgagtagc	tgggattaca	ggtgtgcacc	5940
atgacaccct	gctaatatgt	tttgtatttt	tagtagagat	ggggtttcac	cgtgttggcc	6000
aggctggctc	tgaactactg	acctcaagtg	atcacccgc	ctcggttcc	caaagtgtctg	6060
ggattacagg	cgtgagccac	cacgcatggc	ctattttata	ttattttaga	gtgtacgcct	6120
tatacttata	aaaaaaagct	aactgtcaaa	cagcctcggg	caggtccttc	aacagatatt	6180
ccagaagaca	ttgttatcat	aggagatgac	agctccgtgc	atattattgt	ccctgaaaac	6240
cttctagtggt	ggaagtggaa	gacagtgata	ttgatgatag	gacccagtgt	aggcctaggg	6300
taatgtgtgt	gtttgtgtct	ttgcttttaa	caagaaagtg	taaaaagtta	aaataaaaata	6360
caaaaattttt	taaaatagaaa	aaagctgccc	aggaacaatg	gctcacacct	gtaatccac	6420
cattcgggga	ggccaagggtg	ggtggattgc	ttgagctcag	gagttcaaga	ccagcctggg	6480
caacatgggtg	aaaccccatc	tctacaaaaa	atacaaaaat	tagccgggtg	tggtggcatg	6540
cggctatagt	tccagctaatt	cgaggggctg	aggtgggagg	atcactgggg	gggaggtggt	6600
tgaggctgna	gtgagctgtg	attgnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6720
nnnnnatattc	ttaaaaaaat	ttttttttat	ttttgagaca	gaatttctct	cttgttgccc	6780
aggctggagt	gcaatggcgc	tatctcagct	cagggcaacc	tccacctcct	gggttcaagc	6840
gattctcctg	ccttagcctc	ccaggtagac	gcgcccgcga	ccatgctcgg	ctaatttttg	6900
tatttttagt	agagatgggg	tttcaccatg	ttgtccaggc	tggtcttgaa	atcctgcctc	6960
aggtgatcca	ccccctcgg	cctcccaaag	tgctggaatt	tacaggcgtg	agccactgtg	7020
cctggcctcc	tttacatttt	tttaaattta	atttttaatt	tttaattttt	aatttctcat	7080
atataatat	ttttaagact	agccaagtga	agcagtggga	gtggaaaagg	aactggtttt	7140
gatcaatagg	tgtaaacacc	actgcactgg	gaccagccta	ttttacattc	ctgttagcag	7200
tgatgaggg	tcacttttct	tgtagcctca	acaattatgt	tcgttgccca	tctttttttt	7260
tttttttttt	tttttttttg	agatggagtc	tcactctgtt	gcctaggctg	gaatgcaatg	7320
gcatgatctc	agctcactgc	aacctccgcc	tcccagggtc	aagtigattct	tgtgtctcag	7380
cctcctgagt	agatgggatt	acaggcgtec	accaccacgc	ccggctaatt	ttttgtattt	7440
tcagtagaga	tggggtttca	ccatgttggc	caggttggtt	tcgaactcct	gacctcaagt	7500
gatecgccca	cctcggcctc	ccaaagtgtc	gggattacag	gcatgagcca	ccgcgcccg	7560
cctgcccato	ttttttttgt	tatagccatc	ctagtggatg	taaagttttt	ttgtgatttt	7620
gatttgtgtt	tccttaactga	tcaatgatgt	tgagcatctt	ttcctgtgct	tattggcttt	7680
tggtatatct	ttggagaaag	gtctattcag	gtcctttgcc	cactttaaaa	ttaggttatc	7740
tttctattac	tgagatgtaa	gagttcttta	tgttctagat	ataagtctcc	tacatatgat	7800
ttgtaaaaat	tttccttcca	ttattgggtt	gtcttttact	ttccttttgt	gtcctttagt	7860
gcacaacagt	ttttaatatt	gaagtccaat	tttctatttt	tctcttttgc	cacttgatc	7920
ttggtgtcat	gtttaaggaa	ctattgccta	atctcaggtc	acaaagattt	acacctgtgt	7980
ttccttcttt	ccttcccttc	ttccttccct	ccttctttcc	ctccctccct	ctctccctcc	8040
ctccctctct	ccctccctcc	ctccttccct	tcctccctcc	ctccctcctt	ccttccctcc	8100
ttccttccct	ccttccctcc	ttccttccct	ccttcccttg	tccttctgac	ggaatcctgc	8160
tctgtcaacc	aggctggagt	gtagtggcac	gatcttggct	cactgcaacc	ctgcctcct	8220
gggttcaage	aattctcctg	cctcagcctc	ctgagttagt	gggactacag	gcacacacca	8280
ccatgccccag	ctaatttttg	tatttttagt	agagacgggg	tttcaccaca	ttggccagga	8340
tgggttccgat	ctcctgacct	cgtgatccac	ccgccttggc	ctcccaaagt	gctgggattg	8400
caggtgtgag	ccaccatgcc	cggcctgtgt	tttcttagag	ttttgtagtt	ttagctctta	8460
tagtttagatc	cttgatccat	tttgagtga	ttttgatata	agtgtgagat	atccacctgg	8520
tgttgtaaat	tgcccagaag	tgggtatgct	tctaaatctg	gctgttaggg	attactagag	8580

gtgaccaaag	tgaatttttt	ctttgtttct	tttttttttt	ggagacagag	tctccgtcac	8640
ccaggctgga	gtgcaatggc	ttcatcttgg	ctcagtgcaa	cctctgcctt	ctgggtttcaa	8700
gcagttctcc	tgcctcagac	tcttgagtag	ctgggtattac	aggcgtgtac	caccatgctt	8760
ggctaatttt	tgtattttta	gtaaagatgc	agtttcacct	gttgccagg	cttttctgga	8820
actcccggcc	tcaagtgatc	catctgcctc	tacctcccaa	agtgcctgga	ttacaggtgg	8880
gagccaccgt	gcccagtcct	tttctcagaa	tttatttgtt	tttttttgtt	ttgtttcatt	8940
tttgagatag	ggctctcactc	tgctcagctag	gcaggagtgc	agtgggtgta	tcattgctgc	9000
agccttgaac	ttctggactc	acgtgatctt	cccacctcag	cctcctgagt	agctaggatt	9060
acaggcatgt	gcttccacac	ctggctaatt	ttttaatttt	ctaggactta	tttgtccatt	9120
cttgcaaagc	aggggtacaac	atgcctatct	ctacctacct	ctcttcctt	caagggactc	9180
cagccaaaat	ccttgagggt	ctcgggctga	ctgtgggtgc	tggtgcctga	tctgcctcag	9240
tcatgctgca	tgatcaaaaag	tgcccgtttt	ctgcttcttg	gaactttatt	cactttgggt	9300
gtcagctcttc	ctctgcagtg	tcccaagaac	acagaattag	accaggaatc	tgtgttgcca	9360
tagtgtgtgg	aaagaggcag	acttccaact	ccgctatgtg	ctgttgggtg	attgaagctt	9420
aattttcttt	ctatctttct	ttcttttctt	ttcttttttt	ttttttggag	atggaatctc	9480
gctctgttgc	ccaggctgga	gtgcagtggt	gcgatctcac	ctcactgcaa	cctccgcctc	9540
ccaggttcaa	gcgattctcc	tgccctcagcc	tctgagtag	ctgggattac	aggtgcatgc	9600
caccatgccc	ggctaatttg	tgtaatttta	gtagaaacag	tgtttcacca	tattgggtcag	9660
gctggctctcg	acctcctcac	ctcaggtgat	ccaccgcct	tgccctccca	aagtgtcggg	9720
attacaggcg	tgagccaccg	tgccctggcac	ttaattttct	taataacctca	attaccccat	9780
atggtaaaaat	gggactagta	atccatacct	tatagcgctg	ttgtgaaaat	gaaatgaggg	9840
taagcagata	aaatttcaga	ctacggatgg	gattgttact	acattctgaa	cctggctttg	9900
ctgttatttg	ctatgtgacc	ttatcttctc	tggatctcca	ttctttccaa	gtctataaaa	9960
caaagtggac	aattgtcaac	ctttcttcca	aagagcaatg	atttaaggat	caaagtatgt	10020
cattttaacaa	aaatatgaag	agctcaacaa	atgaggaaat	cattattatt	attacaatta	10080
ttattatttt	agaaataggg	tcttgttctc	ttgcctaggc	tgaggtccag	tggtataaac	10140
acagctcaat	gcattcttcag	cctcctggat	acaagtgatc	ctcatgtctc	atccccctaa	10200
gtagctggga	ccacaggcat	gtaccaccac	gcacggctaa	tttttttatt	tttattttta	10260
tttttttgaga	cagtcttgot	ttgtcgccca	gactggagtg	cagcagcgca	atcaccgctc	10320
actgcaacct	ccgcctcctg	ggttcaagtg	attctgctgc	ctcaacctcc	caagtagctg	10380
ggattacagg	cctgtgcccac	catgcccggc	taattttttt	gtattttttg	taaagacggg	10440
gtttcaccat	gttgcccagg	ctgatctaga	acccctggcc	tcaagtgatc	cccctttctt	10500
ggcctcctaa	agtgcctagg	ttacaggcgt	gagcctctgc	acctggcctc	ggctaatttt	10560
ttattttttg	tagagacagg	ttctcactat	gttgccaggg	ctggctctga	actcctgggc	10620
tcaagtgatc	ttcccacctc	agcctcccaa	agtgcctgaga	ttacagatgt	gagccactgt	10680
gcctggcctg	gaactcatta	ttgaagcatt	cactagtatc	aactttgggg	ttacctggcc	10740
acatcctctg	acctacctat	aagggtatca	cagctaaccg	agcctctgtt	tctcagaatt	10800
taggcagaag	cagttcaatt	tatcacaaac	tactctatat	ccagcataag	tgcccaaata	10860
aaacaattgc	taaagtctct	taggcattta	ctgtttgtta	gttagatatt	tagtcctcac	10920
tacaaatctg	tgatacaggt	attattttta	ttaaacccat	tttatagaag	agaaacctga	10980
agctcagaga	tgctaagtaa	cttggtgcaag	gtcacacagc	tagtaaataa	agggcagagt	11040
aaagattttg	tttcacattg	gactccagaa	cctttctact	gggactcatg	ggaatagtgt	11100
ggatgtccct	gaccttcagt	ggcccagggc	tctcctgggg	gaatccagcc	atagacaaga	11160
caccagcgag	agcccaatcc	taagattttg	tttgtttgtt	tttgagacaa	ggctctcactc	11220
tgtcaccaga	ctggagtgc	gtggcatgat	caatgctcac	tgcaaccttg	atctcccagg	11280
ctcaagcaat	cctcccacct	cagcctcctg	agtagcttgg	actacaggtg	cacaccacca	11340
cacctgacta	atttttaaat	tttatttaat	taattactta	ctattatttt	ttgagacagg	11400
gtatcacttt	gtcacccaag	ctggactgca	atgggtgtgt	ctcagctcat	tgcgctcctcc	11460
acctcccagg	ttcaagtgat	cctcccacct	cagcctctgg	agttgcaggg	actgcagggtg	11520
tgccccaacta	tgctcagcta	atgtttttat	tttttgtata	gatggggctc	cactatgttg	11580
ccagggttag	tctcaaactc	ttggactcaa	gcgatcctcc	tgtcttggcc	tcccaaagtg	11640
ccgggattac	aggcataaac	caccacaccc	aacccttaag	gtgtttttgc	tgaatgtgac	11700
catgtcagag	gcaggaaagg	gaagcatcat	ggggtaggga	aaggaaact	gagcagggag	11760
acaaagaaaa	tgggatcatt	ttgtgagtgt	tcgctgtgtg	tgtatgtgtg	acaattctca	11820
gagccagcct	ctcaggtgg	tgagaccaca	gtccccattt	cccagatgag	ataatggagc	11880
ctcagagagt	ttctgcagca	cagctagtgg	aattagaatt	tgaacccggc	tcttccagac	11940
tccaggtgct	tcacaacccat	cccaaaccta	gtcattttgca	gtttaccttc	atgattttac	12000
catttccctt	tgccatagct	agtgttattt	acttaataat	tccttttgaa	tcagtctgct	12060
taaaaaaaaa	tagcttcatt	ctaaagtgt	atattcttgg	aatatcgggt	ttgctgttac	12120
ccacccccac	acgttatata	tatacatgta	tgtttcta	acatatatat	gtacgtatat	12180
acgtgtatcg	ttttttgtta	ttttttttgt	tggtgttagt	tttttttaga	tggagtctct	12240
ctctgtagcc	caggctggag	tgcaagtggg	tgatttcggc	tcaactggaac	ctctgcctcc	12300
tgggttcaag	cgattctcct	gcctcagcct	ctggagtagc	tgggattaca	ggcaccacc	12360
actacacccg	gctaattgtt	gtatttttag	tagagacagg	gtttcaccat	gttgccagg	12420

tgggtcttga	actcctgata	tcaagtgate	cacctgtttt	ggcttcccaa	agtgtctggga	12480
ttataggtgc	gagctactgc	ggctggccaa	tgtatgtttt	taatacacat	tcaaataacg	12540
aataactatg	aaacctgaaa	aactgctcca	tgttacttcc	tgaacccatc	ttgagtgtct	12600
acatgtctgt	cataccacat	attgggaaac	actgttttcc	ctggcttcca	agcccagctt	12660
aatcactgtc	ccatccctatg	cttcgcttta	tttgtctata	aatgttgggg	ttgggggttg	12720
atgccaaaga	cctttttctgt	tgtcattaac	atggacacag	ctctaagagg	tcttggcatc	12780
ttgggctggc	tctcctttta	gttcagaatt	tggattttta	tccaactact	cagagtgate	12840
aagccttcct	tatgaatgaa	ctcgttggtc	aaactcataa	aaggctgate	gataaaacag	12900
gaatgaatgt	atgaattgac	actaagtcac	tagcattttca	cgggaatgga	ttctccgtta	12960
gtggaagagc	acatgtccct	tctggcactg	atgtgtgctt	gggaaactta	ctgagctaac	13020
tggcccatgt	aacacagagg	ccctttgggtg	cagtggaaaa	ctgttgactt	tggagattat	13080
cttgagtttg	aatctgagcc	tgcctgtaag	aagctggcta	actgaattgc	tttgccttct	13140
ggacccttac	catttataaa	atggggacca	ttgtactcac	cctttagggt	tattgcatgg	13200
attaaatggg	atttctctata	gaaaatattg	gcacaaagta	ggtgtaaatt	tgacgcctag	13260
tgggattgtt	tgtgagggaa	attgtcattt	gattatcaaa	gacttaggag	caggaacagt	13320
gtctaattca	gggactgcaa	atggaaatgc	cagctgaggc	caggcaattg	ctaataattg	13380
ggtaaagcag	ggcaggtgta	gaatagcaat	gtctgggaat	taaaagagag	gtgaggacgt	13440
gtatgacctt	gagaaggcaa	gccctggcaa	aaggggatgg	cctccactca	gctacagtca	13500
tgcctagatc	ttctaacttt	ttatttttat	ttttattttt	tgagacggag	tcttgctctg	13560
tcaccagggc	tggagtgcag	tggcgcgata	tccgctcact	gcaagctccg	cctcccgggt	13620
tcacgccatt	ctcctgcctc	agcctcccaa	gtagctggga	ctacgggcgc	ccaccaccat	13680
gcccggctaa	tttttttttt	gtatttttag	tagagatggg	gtttcaccgt	gttagccagg	13740
atggtctcga	tctcctgact	ttgtgattta	ccctccttgg	cctcccaaag	tgctgggatt	13800
acaggcttga	gccaccgcac	ctggccgata	ttctaacttt	ttaaagagaa	gcaagacatc	13860
tggattttta	tgtgataact	cctgatttta	aactggcacc	caattataat	ttacaacact	13920
ataagggtca	acattgccag	cagagcaaaa	catgggtggg	ggcaactgct	ggtcaccggt	13980
gtgcagcctc	tgggtctaaaa	tcatctttgt	atttcttctt	gctttacgca	ttgtcccagc	14040
acagtgtctg	tgtatagtaa	atatccagta	agtgggtgta	gaatgaataa	accaatgcag	14100
ataaacctgt	agagaggccg	ggcacagttg	ctcatgtctg	taatctcagc	acnnnnnnnn	14160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14280
gggtagatca	cctgaggtca	ggagtccaag	accagcctgg	gcaactttgg	aggccaaggt	14340
ctctacaaaa	ataaaaaaat	tatctgggca	tgattgcagg	tgcccttaat	gaaaccccgt	14400
ctctacaaaa	ataaaaaaat	attgcttgaa	cctgggaggg	ggaggttgta	gtgagccgag	14460
atcatgccat	tgactccag	cctaggtgac	ggagcaagat	tctgtctcaa	aaaaaaaaaa	14520
aaaaaaaaaag	aaaaaaagaa	agaaaaagaa	acaatgaatg	agtgtgaggc	tcatggtagt	14580
attggttcct	gagagttagc	aaccttattg	gtcatcccag	ccacgaagtg	aaatggtacc	14640
cctggcttgg	gccaatgaat	gaggaagaat	aatggcaaat	gggggtctat	gcctccaccc	14700
tccaccacta	gggaggtctc	aagcttgaaa	tccagtgaac	agggttttag	gtcctggacc	14760
tggccagtc	tctacagtc	aagtagataa	gtggaggggt	tgggtccgtg	ggctacggag	14820
atagtgatca	aggccgttac	tctgcaatca	gactcagaaa	tggcctctca	gttacttctc	14880
catttgtggg	tcttttgtaa	gagcagagaa	gaggaaggaa	tttaggtctt	ctcacctctc	14940
gggtgcctg	tccctgctcc	ctgagccatg	gagggctggg	gtggaatatg	gggaataaat	15000
ctgtactttt	tttttttttt	tttttttgaa	cagagtctcg	ctccgtcgcc	caggctggag	15060
tggcgtggcg	tgatctctgc	tcacagcagc	atctgcctcc	cgggttcaag	ttattcttcc	15120
acctcagctc	cctgagtagc	tgggattaca	ggtgccacc	accacgcccg	gctaattttt	15180
gtatttttag	tagagacagg	gtttcactgt	gttgggcagg	ctgggtctca	atacctgacc	15240
tcaggtgatc	caccgcgaca	tgccctccaa	agtgcgtgaa	ttacaggcat	gagccaccgt	15300
gcccggctct	accaatctgc	acattttaat	tgacaagggt	caccctccac	tcatgtgcca	15360
ggcatagttc	tgagaagcat	cccacaagga	tgccctctgag	ttcaccctga	caagtccact	15420
agctcttggc	agagacatct	ggcaaatcca	aggcttgaga	catgctggcc	tctctttaaa	15480
gtgcagcaaa	ttttgtctag	agcttgggtc	gttaaaattt	tgatgttttg	ttttgcatta	15540
atttcaattt	ttaagaaatg	ttgcattaaa	atgttattta	tcttgaatag	taaatttctt	15600
agtgtccctt	taatttctta	gtgtgtctga	gttgagagcc	tcccctgcct	gattctagtc	15660
cagaccctgg	ggtgacagaa	gactggtggg	agatgggagg	tgaggagggg	agtgttggtt	15720
ggagaggatg	atctacagag	tgctggagag	actctgtatg	gagcttttca	tgctgcctgt	15780
ttgccagccc	tgaagctatg	ccttgagggt	gggcaagggt	gcatactcta	gatcagagat	15840
cctcaactgg	ggccattttt	ctccccagag	gacatttgga	aacatgtgga	gacatttttg	15900
atcatctgcg	gggttgggga	gaggggctac	tgacatctgg	tgagttagaga	ccagagggac	15960
cattaaactt	tctacaacgc	ccaggacagc	ccctccacaa	taaagagtta	tttgacctca	16020
catattaata	gcacaaagtt	gaggaaacct	gatctagatc	cacagcacag	aagaaaggat	16080
gtagattttt	cacacattaa	agatgagaaa	gcttgtgcct	gtaatccctg	tgactcagga	16140
ggctgtggca	ggaggattac	ttgagcccag	gaattcaggg	ttacagtga	ctatcatcgc	16200
agcactgcac	tccagcctgg	gtgacagagc	aagattttgt	ctcttaaaaa	aaaaaaagat	16260

gaggacaggc	acagtggctc	atgcctgtaa	tccagcatt	ttgggaggcc	gaagtgggtg	16320
gatcacgagg	tcaggagttc	aagaccagcc	tggccagcat	agtgaacccc	catctctact	16380
aaaaatacaa	aaaattagcc	agctacttgg	gaggctgagg	caggagaagc	gcttgaaccc	16440
gggaggtgga	gcttgcaagt	agccaaaatc	ttgccattgc	actccagcct	gggcgacaga	16500
gcaagactcc	gtctcaaaaa	gaaaaaaaaa	aaagatgaga	aagaggaagg	gagagaaaaa	16560
agagagagag	gaaagaaaga	gagaaggttt	tggagtcaaa	aagacttaga	aattccagtt	16620
cttccacttc	ccatggaacc	ttggcaagtt	gccttctctc	tttctctgaa	tctcacattt	16680
tgcctctgtg	aagttaggggt	ggtacctggg	ggagatgatg	cggagatgag	ggtgaggggt	16740
gtgttgacac	ctatgccctt	aggatgggtg	agagcttggg	agcacgaac	ctccctttcc	16800
cctcttgttt	cttcccccca	ttgtctccca	ccagctccct	gggatctcca	cttcaacttc	16860
tgggattcca	ccagcaggag	gctactcctg	gagttaaggc	gtgttgttca	gactgggggc	16920
tttttaggggg	cataaataat	aattatgcct	ggacaatgga	cataacatct	agggccttct	16980
gaagcaaaac	aggggtgtgg	gtacccaaac	aaggcagtag	gccccaggag	gcaggtccct	17040
gcagtcccag	cagagagcag	ggcacagggt	tgagaagact	gagcaaactt	cattatcagc	17100
tcctttgtcc	cccactctgt	cctggagcaa	tcattctggc	ctcttcccac	ttccccaaaa	17160
accaggtata	aaggctgctt	ctggccccctg	aagccagagg	cactgagagt	ggaggtctca	17220
gactcttgga	aggtgagttc	ttttctggct	gcccaggcag	gaccagtgtg	ggcctggga	17280
agaagcagca	cctcataggg	caaacacgta	ggaggcctgt	ccttaggaac	atcatagcta	17340
agcagacctg	tccccgcagg	ggcaggagtc	tgggctaagg	gtgatactgg	agagcagcaa	17400
cggagactgg	aagacaaatg	aaatttggtg	cctgagttat	ccctcccacc	attccttttc	17460
tagactctcc	agctcagggt	ctgttcatgg	caagaggaga	aagcaatctt	gtttgctctt	17520
taatcaaaac	attaaacaaa	tattccctct	atactatgtg	ccaggggcta	tactagacac	17580
acaaagacag	ccccaaagag	gacggtggag	tagtgtcctc	gctaaaagac	agtagatatg	17640
caatgcctct	tgctcctgcc	ctttctcctg	ctgggaacag	tttctgctct	tcactctggg	17700
aagtctctcc	cttccctcct	catgcgtctt	tccctttttt	cctttttcct	acactccctt	17760
ccccccgctt	ttatttgcac	tcattgagcc	aggaccacag	ccttccctct	ttagctgata	17820
cagctcatct	ccggttaagat	atcacttgga	ctcagaactg	taacctggaa	ctttctcttt	17880
tttgtttgat	ttttttttgt	tgttgttgtt	tttgtttttt	tttttgtttg	ttttttgttt	17940
tgttttgaga	cggagtctcg	ctctgttgcc	caggctggag	tgcagtggcg	cgatctcggc	18000
tcaccacaaa	ctccgcctcc	cgggttcaag	caattcttct	gcctcagcct	cctgagtagc	18060
tgggactaca	ggcacatgcc	accacgcctg	gctaattctt	gtatttttag	tagagatggg	18120
gtttcaccat	atttgccagg	ctggtctcaa	actcctaacc	ttgtgattcg	cccgcctggg	18180
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgcac	ccggcaaaat	gtaacctgaa	18240
ctttcagaag	gaaaaaccac	ccactgttta	agatgaaggg	ctggtgactg	ccccaggctt	18300
ctcacacgtg	ctttctccca	ccttcaaaac	acacactcgt	ggtgtcggcc	agaagtcagg	18360
ttcttgtcca	tttggtgggtg	tgaccogaga	gatctctcct	tacctaacac	caaggaaatc	18420
ctccagtctt	gtcttcagggt	ggaattccta	ggaaagctcg	agcgacgttg	ctggagctgt	18480
ccacggtgct	ggaactagga	agctcttgac	ctgatggcag	gttacctctt	cttcccagag	18540
aatgatgccc	cccatctgga	gagcctagag	acacaggcag	acctaggcca	ggatctggat	18600
agttcaaaag	agcaggagag	agacttggtt	ctgacggagg	aggtgattca	ggcagaggga	18660
gaggaggtca	aggcttctgc	ctgtcaagac	aactttgagg	atgaggaagc	catggagtcg	18720
gaccagctg	ccttagacaa	ggacttccag	tgccccaggg	aagaagacat	tgttgaagtg	18780
cagggaaagtc	caagggtgcaa	gatctgccgc	tacctattgg	tgcggaactc	taaaactttt	18840
gcagaagctc	aggtaaagtag	tagggaggct	actgcggagg	acctggggga	aaagagagta	18900
cattcagctc	tctgttccct	attcatttag	gctagtgggt	ctcaaagcct	cgcatgcac	18960
agaatcacct	ggagttgttg	ttaaaacaca	gctttctggg	cctcacctgc	acgacttctg	19020
atthagagg	gctgaggtga	agcctgagaa	tttgcattta	caacaaatcc	ccaggtgatg	19080
atgatattgt	tggctctgggg	agaaccaccg	atttaaaca	aaggctttgg	tgttagaaac	19140
gcctgtgtta	aattctggtt	ctgcctttta	ttagctgtgt	tacctgggga	agttgctttg	19200
cctttcaaag	ctttagcacc	ttcatttgta	aaacgaagat	atatagcacc	aacttcttag	19260
agttgtggtg	agcattaaat	gagataatac	atgaaaagtg	tttggaatag	tcactgggct	19320
gtaataaact	ctcaataagc	ggtggttata	attattatga	gtattatcat	ttcctgtagg	19380
attgtcctga	cagctaatta	agaagcaaaa	gataggatta	agggaggcaa	gtaggtttat	19440
ttttaacctt	aaaagggatg	ccgggctctt	gcctggagac	tcagaaactt	gaaataaatg	19500
agaggaatt	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19620
gtagacacat	agccagaaca	tctagaaggg	gtggtaggag	tggggattag	aggttccagc	19680
tggaggcaat	ggcacttgca	aaggctttgt	tgaagtggcg	taagtgtgga	ggtggagcat	19740
tcaggaaagg	agagcttcag	cttcagtggt	gctggagtg	tgggtgtgaa	gagaggtgaa	19800
gatgaggctt	ggaggctggg	cagattttgc	tccaaaagag	cttgggtgaa	tgtgataagg	19860
agtttggatt	ttctcctact	aaggacaaca	gcaaactatt	gaagagttaa	aatcgttcag	19920
tgacaatgac	acgtttgcgt	tttggtggct	cactcgagct	gccagccagg	tagacagtgg	19980
cagaagatgg	aagataaagc	actaaagggt	gatgaggcag	gaagccagtg	aggagagaaa	20040
ggggacgatg	tgagtgcacg	taaatcattt	gttgggtgtg	tattgtgtgc	taagctctgt	20100

gctaaattct	tcacgtgtat	tatttcagct	aattccatcta	acaactctgt	aaggcaggta	20160
caatcggtcc	cagctgaaga	agctgaggct	ctcaaaagct	agtaacttgc	ctaagttcat	20220
gcagcatgca	agttgtccag	ccaggattct	aacttagaca	ccagaggcca	cttttaacca	20280
ctgctctagg	actgggggaa	atgggtcccta	gtgagatatg	tgtcagagttt	catatttcat	20340
tcaacaatat	tgttggcctg	ctacatgtga	agagctgtgg	aaagcgccca	aagtgaagtt	20400
gatccctatg	agcaagtggg	atgggggtgg	agtggacagt	aggagggtg	gaacacacat	20460
aaaagggtat	aagaaataac	aattaggccg	gccagggtg	gtggctcacg	cttttaatcc	20520
cagcactttg	ggaggccgag	gagggtggat	cacttgaggc	caggagtttg	agaccagccc	20580
ggccaacatg	gtaaaacccc	atctctacta	aaaatgcaaa	aattagctgg	gctgggtggtg	20640
cacgcctgta	atcccagcta	cttggggaggc	tgaggcacga	gaatcacttg	aaccagaggag	20700
gcagagggtta	cagtgaactg	agattgcacc	actctactcc	agcctgggag	acagagtgtg	20760
accctgtctc	aaaaaaagaa	aacaaaacaa	gtaggtactt	tctgccatag	ggaggattca	20820
taaactgcta	gtcctcaggt	gcatttttgc	ttatcagttt	taaaaatcag	agaatgtctc	20880
aaagaattag	gatgtcagct	tcttttgaaa	atttgggcca	gaagcgggtg	ctcacgcctg	20940
taatcccagc	actttgggag	gctgaggtgg	gtagatcacc	cgagggtcagg	agttggagac	21000
cagcctgacc	aacatggcga	aaacccgtat	ctactaaaaa	tacaaaaatt	agctgggtg	21060
gtgggtgcatg	ccttttagttc	cagctactca	ggatgctgag	gcatgagaat	cacttgaacc	21120
cgggaggcag	gggttacagt	gaaatgagat	tgcaacctg	cactctagcc	tgggagacag	21180
agcaagaccc	tgccctcgaaa	aaaagaaaaa	gaaaatttgg	aagatctgac	aacagttgac	21240
ctgcattcct	gctcggcaac	agcctgatgg	tggatgggca	gaggctcagt	tgtctgccaa	21300
acctcccatc	actgatgtct	tccctcgctg	tcacatctg	cttgacatgt	aggcatttgg	21360
tgtgtgcctt	ctgctctggg	tgcccagatg	aattggatgc	tatatgagaa	aacattctgt	21420
aaatgtcttg	tggtaggcaa	cctcaaagat	cactggggcc	tccaatgate	cctccttctc	21480
ggtattcatg	cctgtgtata	atcctctccc	ttgagtgtgt	actacacctg	gatacttgct	21540
tctaataaac	agaacacagc	aagggtaatg	ggatgctact	tctaagggtta	aattacaaga	21600
gtgtaaagtc	tgtcctgttt	gtttccctct	cttgatcttc	ctctcattct	ctctctctcc	21660
ctctctctca	ctttcttact	gtcttgtcct	tccctttgtt	tactctgatg	aagcaagcta	21720
gcaagcatcc	atgttgtgag	ctgacctatg	aagaggccca	tgtggtggta	aggaaactgag	21780
ggcagcctct	acccagcaag	gaactgagtc	actcatcata	tgggtgagct	tggagacaaa	21840
tccttcccca	cttgagcttt	cagatgacgg	cagccctggc	tgatgctttg	caggcttgtg	21900
agagaccctg	agacagaaca	ctcagctaag	ctatacccta	tctcctgaga	tagagtataa	21960
tacatgtagt	tttaagctac	tatgttttgg	gataatttgt	tactcagcaa	tagataacca	22020
atacatatac	catgtacata	actgtttcag	ttgtctgaga	ctatatattg	tcattttaca	22080
cctacatcaa	gaatgtgtca	ggcaccattc	cagggtactg	gaatacatca	attaacagaa	22140
taggtaaaga	ggccaggcat	agggctcaca	tctataatcc	cagcactttg	ggaggccag	22200
gtgggaggac	tgcttgagcc	caggagttga	gaccagcctg	ggtaaaaatg	tgagacactg	22260
tctcaactaa	aaaaaaaaaa	aattagttgg	gcacagtggc	acatgcctgt	ggtgccagct	22320
gctcaggagg	ctgaggtggg	aagatcgctt	gagcccgga	gtttgaagct	ccagttagcc	22380
acggtcacaa	aactgcactc	tagcctgagc	aacagaaaaa	gaccctgtct	caattaaana	22440
aaaaaaaaaa	aaaaggaaag	aaagaaaaaa	ataggtaaag	atccttgatt	cttgccctct	22500
tggaaacttct	attctagagg	gggatgggtt	ttcacagtag	aagtctgtgt	tgacagcgct	22560
gtttaaagct	ccttcagcat	ctggggaaaa	ggttnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	23940

caggcagtg	ttttcat	ctctgtg	ttacatcc	cagcttg	ctttctcc	24000
agaaccct	gagatcag	gtacctcc	attttac	tgaggaa	gaggctag	24060
agggaacct	cttgctta	aataaga	gctaata	agtgcgt	gtgcactg	24120
caccttg	tgtttag	tacaacac	ctttgag	gatgcgt	tatcttc	24180
ttgcagtt	ggaaacc	gtacaggg	gcacagtt	gtcattc	caagatca	24240
cagctttc	tggcagcc	cagaacct	gttataag	tacacgct	agtctgt	24300
gggctaga	aggtagag	ggtatatt	atatttat	ctgtataa	aatcaccc	24360
aggcttgg	ttttaaa	acaaacac	ctcatctc	acagtttc	acagtcaga	24420
atcagggag	gactcagc	gctgattc	agtcacag	tctcatga	acatagtc	24480
gctgtcag	agggctgc	tcatctga	ggctgact	ggttggag	tctatgtc	24540
ttcaattac	cccatggc	ctccatag	ctgctcag	cacagcac	gctttccc	24600
gagcaagag	gctaagcg	agagaccc	tatcttct	cacataat	cagacgtag	24660
ataccatca	ttctgtta	ttctattat	ggcacagag	aacctgga	tactgtgg	24720
ggagactgg	caaagcag	gaataacc	aggcaggat	cttgaggg	gtcttgtt	24780
ctggagacca	ccattgag	tttttttt	tttttttt	gagacagt	tgctctgt	24840
cccaggctg	agtgcagt	cacgatct	gctcactg	acctctgc	cccaggtt	24900
agcgatttc	ctgcctcag	ctcccagta	gctgggat	acctggag	cttgaacca	24960
gattctgtg	ctgcttttg	tctttttg	ttcatccaa	cagtcctgt	ttatccta	25020
aggatggg	aaagagact	ggagagaag	aaatccagt	gcctccctc	ctgctagca	25080
agcctggcc	tggcactga	ccttcctct	ctacctctg	ctcctaatt	tgagggtcc	25140
ctagcaggg	ccttctgtc	aggacacat	ggcgcctgt	cctcacocca	cctactgac	25200
ctctctctc	ggctggcct	agtgcctgt	attgtgccg	agagagga	cgctggac	25260
tcaggccaag	ctgctgtcc	caggagggc	tctgcttat	tctagggc	ggacacct	25320
ctgaggact	ctgatgag	acggtgtg	agcttccac	ttccacact	ccttcccat	25380
cttggttct	aaaccttca	gtgtgcat	gaatcact	gtggggg	tttgtccaa	25440
tgcagattt	cagatatcc	cgctgagat	ctgaggccg	agatgagg	tgtgaatct	25500
catgttaaga	aagcacccg	tttgatgct	gtgtcattg	gtaggggag	aacacttga	25560
gaaacatga	gctagaga	gtgggttt	atgggttt	catagaaac	tggatttct	25620
tgttttctg	tgccctgac	tcgaaggac	atctgaagg	ggaggggca	ggccaaga	25680
cagggagtc	tggaacgta	gaggcagca	ccagtga	cccgtact	tcagggacg	25740
ggttgtacc	aaatccat	gaccacgta	ctgcaagat	caaaagct	ccaatgcc	25800
cagtgcctc	acctgacca	acctgtgt	tgtgggttt	ggcatcac	gtctcatca	25860
caacttac	ctccaggta	ccacctcat	ccttccctc	tccctgcct	ccgaggetc	25920
tccaaagg	tggtccat	agcacctgc	ttccaggaa	cgagttctg	gtcttctga	25980
ctggatctat	tttccgggt	ctccaggaa	tgtttctagt	agattgggt	ggcgagggg	26040
tgggaattga	ggcccagtt	gcctcttcg	cctaccctc	cttccctcc	cctccacac	26100
ctctcctaac	ctcttctac	tctcttttg	gttttagtt	gtgaccttg	tcctgcac	26160
catgtttcga	ggtttcttc	actcagcct	tgggagtct	tatgctgca	tgtgagtc	26220
ttgggtgaa	atgccttct	gagctttgc	accgtgat	gagaaccca	gggaagggt	26280
gggagggcc	caggcatcc	ctaatagc	tctctctga	acctctgat	ggcagggag	26340
tcacttctt	aaaggcag	tatcctgt	taattgact	cccctgttg	agtcttcc	26400
tagaggaag	tgaaatac	ggcttgat	cactttggt	ctatgtctg	tgtttgaac	26460
ggcccacaga	atggcctcc	ctccatgcc	acctgaaga	aatttccca	gggcagcc	26520
ttgccttata	attttcct	tcatgttga	cagtcccc	ttgcatct	ctcctggt	26580
cccctgctg	gcgctgtga	gggactct	cctgtgtat	tgatggag	acaggacat	26640
acaataatga	tgacaaaat	acaaccat	tcaagtgt	cgttgggt	ggcagcagg	26700
aggatcctt	accatcact	cctgagtt	gcctcactg	agcggctct	gcagaggca	26760
gctctcttc	cttcatctg	tcaagccag	acctggagt	ttccttgat	tttctctcc	26820
tcacactcca	tgttcactc	atcctcag	cagccagca	cagcttctac	acaccccaa	26880
tctgaccctt	cttgtcac	ccactgtgc	ctctccag	ctagccacca	acatctctg	26940
cctggattat	tgtggcagc	tttagtct	cacatctgc	ctggcccgc	tgtctcag	27000
tatttttaac	acaggggct	cagtcacct	tcaggacata	agtctcttc	catcactct	27060
tggtgtcct	tctcatctg	ctcagagta	aagccaaag	ctttactat	gcctaaaa	27120
ccctgcaag	tctggccca	gcacttca	cccctctag	tccccctct	ccattgttc	27180
ctctgccaca	gccacagt	ttcctagt	tccggaagt	tcaagtgt	tccctgct	27240
gcatctttg	atgtactag	ccctgttt	agaacatt	tctccagat	tctgcaagg	27300
gcccaatct	accttctct	cttcttcag	tcttctct	actgtctct	tctcagtg	27360
gcctccctg	gctgtccat	gtacaatt	aacctccca	ctgcccgt	ctctgcttg	27420
tttttctcag	cgtttatca	taacactct	cctatctct	gcttattgt	tgaccgcc	27480
ctgctccatg	ggaatgcc	ctcctcgat	gcaggaat	gttgacttg	ttgatcgt	27540
tatctccagc	acctagag	gtgcctgg	catagtag	tctcagct	atgtttgt	27600
acagaataca	gtggacagt	ctgcgaggt	aatgccat	ctgttattg	tggaggaag	27660
ggggctcagg	gagtttgag	cacttgcc	tatcacac	acaggaggt	tgagaacca	27720
gctcagtg	cctgaagtg	gagcatttg	cctcaagg	ggggaccaa	gagcccatg	27780

aaagagcccg	aacgcttaag	caccaccctg	cctggccagc	ggggnnnnnn	nnnnnnnnnn	27840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27900
nnnnnnnnnn	nnnnnnnnnn	nnnncccaact	gcgcctggcc	cattactttt	aatggcaaaa	27960
accacaatta	cttttgcacc	cacataaata	gttaccatgg	gctgagcatg	gtggctcagg	28020
cctgcaatcc	cagcactttg	ggaggctgag	ccaggcggat	cacttgaggc	caggagtcca	28080
agaccagcct	ggccaacatg	gtgaaacccc	gtctccacta	aaaaatacaa	aaattagctg	28140
ggtgtggtgg	cgctgacctg	taatcccagc	tattcaggag	gcagaggttg	cagttcactg	28200
aaatcatgcc	actgcactcc	agcctgggag	acagaatgag	actctgtctc	aaaaataaat	28260
aaataaataa	ataaatatbt	accatgtttt	gaccacctgt	tatgtgccaa	ctgtattact	28320
taaaaacacc	catgggaggg	tgggcacagt	ggctcacgcc	tgtaatcgga	cactttggaa	28380
gggcaagcgg	ggaggtatccc	ttaaggccag	gagttcaaaa	ccagcctagg	taacacagta	28440
agccctgtct	ctacaaaaaa	taaaaaaatt	aactgggcat	ggtggtgtgt	gcctgtaacc	28500
ccagctcctc	gggaggcaga	gggagaggtt	cgcttgagcc	cagcagtttt	agggttgcat	28560
gagccaggac	caagacacta	cactccagcc	tgagtgcag	agcaagacac	tgcctctaaa	28620
caaacaaaca	aacaaaagcg	acctgtgggt	aggtaggaac	aggctcatag	tacagatgag	28680
aaagcagagc	ttggaggggt	caagcgattt	gccaaagcga	ggtccaagcc	gaggtctctc	28740
tgaatccaaa	gttaattccg	tctatcatat	caccacagcc	ctctctgccc	cagggagagt	28800
ctctgcccac	tccagccact	cacgtgtaat	tgacttcctc	aggggcagga	aaggcttcga	28860
tgggccagtt	gaggtgagc	ttcagaaaga	taaggcaggg	caggccagac	caggtgaaca	28920
tgatgaccac	gaaggccaca	ccggcatcgt	agatcagctg	tgagaggagg	gggcaggccc	28980
gtgggggaga	ctgcctggcc	ccagacccca	ccaaggtaga	tcccaggcct	cagaggcctt	29040
aaagaagtgc	tcttctcccc	ttgtccttgt	gcccaatttg	cagatgagga	aaccaagacc	29100
agaagttag	agtcagactc	agaagaccca	tcattccttt	ttctttttca	cttgaggccc	29160
cctagagagc	tatgaaatag	tctccacaaa	gcctgaagtt	gctggccact	ggctcaaaat	29220
atctctgaaa	tttccattat	cttaaaaaaa	tacatacatt	tttgccctatg	actccacaaa	29280
cattcatgtt	catgttcgca	caaaaatgtc	catttccatg	tacgtacaaa	ggaaaacttag	29340
tgctctaggt	ttaccgggcc	taatcgtgtt	tatcctgccc	cttcctggca	cattccccag	29400
gggaaaaggg	aaacccagac	tgctcatgct	cagccttttc	tcacctttcc	caggtcctcc	29460
cacgtgcaac	aactgggggg	gttggggaga	gggaggtgca	agtgtctctg	ccaagggctc	29520
tcaaccccag	ggcaggtaag	ttctcaattg	aatgagattc	tgtgcaaatg	tgtcagccct	29580
tcttatggaa	gaagctgatg	caccatctgt	cctcttgtcc	tcccataacc	atctgaccag	29640
gataattaat	gtctgtctct	ccctcaggct	cctgctcaaa	cctttttctc	tgcatcttg	29700
gaccttggtg	ccttttctct	cctaggggca	ggacagagct	tcaaagggcc	acacccccaa	29760
atgtgtggag	gtaagatctg	gctcttcaaa	cactacttca	gttgaaaaga	agggagaact	29820
gccacccctc	catgcctgcc	caccagaaca	actgatggcc	ccccaccca	tgcgtctctc	29880
caaactcctt	tggagacact	gagcaaaaag	accttcttta	gtactctttg	taaagtgcaa	29940
aacggtatgc	agtttggtac	tgcccaccgt	ggaggttgag	gagcatggca	tggctcaaag	30000
ggtcctttga	tatttgacag	aggaatttga	ggccccatc	ttgcaactgag	ctaaaacttt	30060
gggtcccttg	cttcgaggta	caccaggttg	acctgtccag	gatccagcct	ggcataaaact	30120
cactttgtga	ccttggaacca	aaccacccat	cctctctgga	aggtgtggaa	aaatgtggcc	30180
ccaaaggctg	aataaagcca	gagagttag	gaccttgaac	gcatgtgaag	gggctggact	30240
tgattctgta	ggtgaagcta	aaccactgaa	ggtttttcag	cagtgtgtga	gccagttccc	30300
catctgagat	ctttctggaa	gtcacgtgag	tgacagagta	cagagaaaaa	gaatcagagg	30360
cagggagacc	agctgagaaa	gcttgctgtg	gccagggaga	gagggggaag	gcctgcattg	30420
ggatgatgac	agagaaagga	gagcggagaa	gtcagacccg	tgggtcagca	ctagctgctg	30480
ctcactcggc	cccacccggg	tcttgtgtca	agacaaaaag	aaaacccagg	tggcctcata	30540
ccttgattcc	tgggaacgta	atggcagaag	aggcgtaaga	gccaatcatg	agggccatta	30600
acgtggagcg	caggttccca	aacatgttgg	gcagctgagg	agggaaagca	gcaacccatga	30660
ggtggggaca	ccgtgaccct	tgcccagcat	tcccagccct	gctccataca	atagctccag	30720
gagacgcagc	agaaaagccc	caaggtaaaa	caaacagaaa	aatcaatgtg	ggaaactgta	30780
ctctgcccc	tgcctacaca	gtcacagtgc	cctttagctt	caaaaaggct	cccagacacc	30840
cctcagagag	acattttgtt	aattttgttt	aattccagggt	ttcccaagtt	tgttacgtaa	30900
cacctctgaa	aaacacatgg	gcactttggg	taagaaaacac	tgatcttggc	tgggcgcagt	30960
ggctcatgcc	tgtaatccca	gcactttggg	aagccgaagc	tggtgggaag	cttgagggtca	31020
ggagtccaag	accagcctgg	acaacatggt	gaaaccccat	ctccacccaa	aatacaaaaa	31080
ttagctaggc	atggtggcat	gcgcctgtaa	tcccacctac	tccagaggct	gaggcaggag	31140
aatcgcttga	acctgggagg	tggaggttgc	agtgcagcca	gatcgaccca	ctgcacttta	31200
gcctgggtga	cagagcgaga	ctatgtcccc	accccccaaa	aaaaaagaaa	agaaaagaaa	31260
gaaacagtga	tcttgtccaa	ccattttgag	atgagacaat	tgagacccag	ggaggaaaag	31320
tgtactcaag	ttcacagagc	acattaatgg	ctttctcccc	attgtcgttg	tcccagccct	31380
aacccaaggc	tgtgaccatg	gctgtgtccc	ggtaataggg	agtgcctctt	aaccctctcg	31440
gttgacgtcc	cagcccagtt	tctgcctaata	caggacaaat	cacatcctgg	gaggtgaggg	31500
tggaaataag	ggagggaact	gagccagggc	agacagtctc	cagaggaggt	ggctctgacg	31560
cagagcaggg	tcagaaccca	caccaggaga	gaatttaatt	gatcatgtgt	tccactcacc	31620

tgccctcagcc	aagccctcag	ggcaggggaa	ggcaaagtca	ggatgccctt	cgcacacacc	31680
ctccctctggc	cccaccatcc	ttcccaagtc	actagatccc	acagctgaga	aggaccttag	31740
gatccgtaca	aagcctaaac	acactccaca	gagggggaaa	ctgagactct	gaagggaggc	31800
ctcaacagct	ctggtaaaaa	aggcgtttag	gccggggcgca	gtgggtcaca	cctgtaatcc	31860
cagcactttg	ggaggccgag	gcgggtggat	tgccctgagct	caggagtctg	cgaccagcct	31920
gagcaacacg	gtgaaacccc	gtctccacta	aaatacgaaa	aaattagccg	ggcgtggagg	31980
cgtgcacctg	tagtcccagc	tactcgggag	gctgaggcag	gagaattgct	tgaacctagg	32040
aggcagaggt	tgcaagtgagc	cgagatcgcg	ccactgcact	ccagcctggg	cgacactgcg	32100
agactccgtc	tcaacaacaa	aaaaaaaaaa	atgggtgttta	aacacataa	actaaattat	32160
ccttccccct	tccccgaag	tggtgtggctc	aggaaaaacc	tctaccact	caggcagagg	32220
ttttcctgca	ccctgcatcc	gtgaggcacc	actggccaag	acgccaggga	aggctgccag	32280
gccctggagag	gggcaggggc	ccctccccct	caagggggcca	caaacgctgt	ctgcgccag	32340
taccgtgggt	aaggcgaggc	cggccggcta	accccgggct	ggcggccttg	cagcgtgcgt	32400
ggcaacagca	gctggggccg	caagactcag	cacgggacgt	cctcgtccaa	gtctggggcca	32460
agagcagcgg	cccagggggc	ggggccggcc	agagggagcg	gggagaggct	gagggggcgt	32520
gccagcgccg	gaccctgcca	ttggctggag	attacaggag	gcggggacat	agcagggagg	32580
agccgctgga	caagccccc	cgggcccga	gggaggtct	gaggtcaaga	gccggagaga	32640
agggatttag	ggccctgggc	caagttgcac	agcagggaga	aggggctgcg	cagaggggcg	32700
gggagaaagg	gatccgcttc	cttcccttag	agctgtgaaa	tgtccccggt	tggaaattaaa	32760
ggcggctgct	ggggagaggt	gaaattcagc	caaaaaccacc	cagtcaggca	gcccttctca	32820
gagataaaca	gtccgagcca	gcccggccag	gaaccttccc	ctccaacctc	cctaagcctt	32880
taacactcct	aagcctttta	cgcgtttaca	cactcacata	aataaacaca	ctttgagcaa	32940
cacacataca	ccactcacca	catgtaatat	gtcaagccat	gtgcacgacg	aggtgtcgac	33000
aatttcatat	ggttcaacct	agtacactca	caaacacacc	taccaactca	tggctttcac	33060
agggacgggg	tcacacaccc	actctcccac	gacatggcaa	gcgtgcacac	gctatctcaa	33120
gctgctccct	ccccctcaag	atcatgttac	ccagttttat	tttcttccca	gcacctaaga	33180
cgactgacat	aattttattag	tttacttggt	tattgggtta	tctgtgcccc	tcacccccaa	33240
aatgtaacct	ccagcaggga	ggatgactcg	gtcagtcctg	attgtgctgt	agtccaggac	33300
ctagaacaga	gtcccatgga	cattcatggg	ctctgtacac	acaacacac	acattaacat	33360
acaccccgac	acacagcctc	atccacacac	acacagcctc	acacctgctc	tttgagcca	33420
cctgcacagt	ttctcacaca	ctcacttgat	ctagtgatct	gcgtccacag	gccccctccc	33480
cagcccactc	atactgcctt	caccccactc	actctgcctt	caccccactc	gggggaactc	33540
tgctgccagg	ccaggcctgt	gacactcacc	gtgagtgaag	tgaacgttag	gcagatgcca	33600
ccaaagccat	tcagggacag	cgccagggaat	atcaacggag	acagagctgg	aaaggggaaa	33660
gcagcagatg	agggcatttg	gggagctgtg	ggaagccaag	ggcgggagct	ggggtaaaaa	33720
tccgccttca	tcccacctat	tcttttcttg	tggggccaca	agaggacaga	caactcacct	33780
tccacgtccc	gggaggccag	ggccatgagg	gtgcaggacg	cagtgaagca	ggcactgtgg	33840
agacacaggg	aagggcgagg	ggttgccctg	tgagcacccc	ccctccccct	ccccctgcagc	33900
acggtccctg	tccctccgtt	ccccatagcc	cagccacctc	acctgccaac	cagccgcacg	33960
ggtcgggggc	caaagcggtc	catgaggatc	cccagtgcca	gggtgggtgg	gctgagcacg	34020
aaggaaacca	tggtgaagcc	caggttgagc	atctcgctct	gctggtcaca	gcctggccac	34080
ctgcgctget	catcctgggt	ggtgttggtg	ctgctctcag	ctgaggaggg	ggaaggaggg	34140
gctcagcaca	tgacaccagg	aacagctggg	cacaggagac	agcagcccac	agtcaggcgg	34200
cctgtcttca	aatcccatgc	caagtgcctt	tgggggtacc	ctagagtcac	atctcctctg	34260
atggggctgc	tccagaaatg	gcagccatta	gtacctgacc	ctgggagagt	cttgtgcaca	34320
cacagcctga	ggcttcaact	agctcaaatg	aaatactgga	cataaaagta	tttactaagt	34380
tgtaatatgc	actcagtgtc	caagcttagg	gggtgtgga	cccccaacaa	gaagtgcctc	34440
catatctaga	ggcaaaggca	aaggcagtga	gtggtactct	aatggctata	acaagaattc	34500
attaaaatgg	cccggcggtg	tagttcatgc	ctgtaatccc	accactgtgg	gaggctgaga	34560
caggcagatc	gcttaagcct	acaagtttga	gaccagcctg	ggcaacatgg	taaaacccca	34620
tctctaaata	aaaaaaagaa	atttagaaag	aacactaaaa	cttagaggaa	gctttcccga	34680
taaatgatag	tctgataaaa	taatagctaa	tacttattga	gcacttaact	atgctccagg	34740
cactgttata	agtcagttaa	ttaaaglatcc	cgttccctag	gtgatgaagc	tgaggcacag	34800
aatgagaacc	aggcactgcc	ctccagtccc	ctctagaagt	ccacttggag	gacttgtcct	34860
taacggtaaa	ctgccaaact	ggagttgtga	caagtttaagg	agaaaagcta	gtgataggag	34920
acaaagggct	gcttcgcttt	actcaatgct	cannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35040
nnnnnnnnnn	nnntccctccc	tccctccctc	cctccctctc	ttccttccct	ccttccctcc	35100
ttccttccct	ccctccatct	tctccacacc	tggtatcatc	atacagaagc	agagaggact	35160
gcacttggtg	aaagtttcaa	ttctcctgtg	tggagagggtg	agcactgagg	aaggggtggg	35220
ggctgtcaaa	ggagacttac	ccaatctttc	cagccaccca	atcccttgcc	cagtgtttct	35280
ataaaaataag	ggccttttgc	atctgattta	agtaggaagc	tgattcctga	gcccctcaga	35340
tctgctgaat	cagatagcta	agggggccct	ggaatctgca	tttttagcaag	cggaggtggg	35400
ttatgaagca	ctccaaagtt	tgagacacgc	ctcaaagggtg	gagtggttct	gtgggggggca	35460

gaaaggaaaa	tgcaaagggg	gaaggggtca	cacttgggga	aggtttcaga	caataccgag	35520
tggaaagggg	gatgccaggt	gtggggagta	acagatagag	gaggcaaaat	gagtggagac	35580
caagccagac	cggggaggag	ggggccacag	caaaggtgag	acaggtcagc	agccagaaac	35640
cgaagcagac	acttgcaggg	tgacccccgc	cctctcttcg	tggcaatctg	agaccgagga	35700
cgtggagacc	ctggagagcc	cccaaccttg	tttctggggg	gtgggtcaga	gaggaagcct	35760
ctcatccccg	ggcaccagcg	gccttcccg	gaggctcaac	acgcagatac	ctggataggc	35820
gtccatcact	ccccggcca	gagcccacca	acgctcctcg	aggtccgacc	ttgtccctcc	35880
ttctacccca	cagtccccag	tcctagctca	tctgcataaa	gctccaatta	acatgttttt	35940
cctttgctat	ttgcgatccc	agaactcggt	ccccaccccg	agccccgttc	ccgccgcttc	36000
ctcgccccctg	ggagggcggc	cccattaacc	ctcgcgaccc	gggcccgtcc	tggcggtcct	36060
gaccccgcca	ccccgtcccc	cggcggggt	ctgggggtga	ggggcgcgcc	ctggggcaga	36120
ggattgcgcg	gcagggtctg	ccacagggca	gaggccaggg	ctctccggga	aaaaggcagg	36180
cgcataatag	cccccttttc	tgggaaaaga	cggggagggg	ggcttctcct	gggagactcc	36240
aggcttcgaa	attcctcggt	ccctatcctc	cggcccccg	acccctcctc	ctccccgcca	36300
cgcaccctct	ccccctccca	gccatctggt	ccactccgca	gcgcgcgcac	aaacacggct	36360
ccagctcgct	tccgccccctg	cccagccccc	tccccaaagc	ccggggagtg	ggggagtgag	36420
cagacgccct	tctcctagga	ggccggaatt	tctgcctcca	tctcccaccg	gggtccggct	36480
ggccagaggc	aagcttcgag	acccccacc	aaccaccacc	accgttgcca	gggcccgtga	36540
ggctgcagat	aacgcttgca	aggacgggag	tccggggagg	tgtaggcgca	gtttaaagga	36600
cgggcagagc	aagccccggg	aagaggcagg	ggttttccct	cccgggtcgc	cgccccccgc	36660
accctcggag	ccagccgcag	ccacgcagcg	cgcctcccg	ggcacaccaa	ggacctggcg	36720
cgcagctggc	gcttaccccc	acccccgggt	cgcctcctgg	ctcgcgctca	gcctccccag	36780
actattcgca	aattgaggat	cccggacaca	gagtgcagag	accccgga	gcctactgaa	36840
agccagccga	acccgctggt	gggtgctagc	caattctgat	tttgtacttt	acaaaaacaa	36900
aaaaagtcag	tgttggaagt	cgggagtctg	ggctcagagc	agcagggatc	tgcatgtga	36960
ctttgccaag	tctccagacc	cctgaggaca	ggttttccta	tctgaaaacg	gagggagacg	37020
tctctcttat	taactttcta	agagaaacaa	agacaaaagg	agggaaaatg	gcttagctgg	37080
aatgctgtct	tacagagcca	acctttggag	gtgggggaga	tggccaaggc	ctctgaggtc	37140
actcttgccc	ccaggagcag	ctgagaaccg	gaaagaagct	tgggacctcc	tttctgcaga	37200
gctatccttt	ccacagactg	cggaggttcc	aaattgagct	ccaccacct	acactgtgtg	37260
cccttgggtg	tgtgccttaa	cctctctggg	cttgtttcct	acagcgacaa	gaaagaatga	37320
caacaccaac	ctcttaggct	atagtttgga	taaaatgaga	tagctgtgtg	gaacagacag	37380
atcctaacc	aatgttagtt	tcccttct	ttggggactt	gctctaacct	ccagggtcta	37440
tgtcccagag	gcacaagcag	gtgcagggt	ggataaataa	ggtagtctt	tctgcaggat	37500
ctcttgctct	cactgatggt	gtcttctctt	gatataagata	attttaagc	ttcacgttat	37560
ttatttat	actttaaagc	ctcactttta	tgttaaaggt	aatgtataat	atagtataac	37620
aaggaagctc	aaaatttgca	taaagtttta	agataaaaata	ggagactcca	aaaaagtgtt	37680
actttcgcca	ggccttaggg	atgctatggt	gggaagtttg	agtcatacct	tagcattctt	37740
tctaaagcat	tctgtcctaa	tcctctgtat	ggagaaaagc	cagcttccctg	gatgtacccc	37800
aaatcctggg	aagtaggggg	caggagctgg	actccctcca	agcactaagg	gcagggcag	37860
ggtgggaaca	gggaggtgag	ccagacagcc	agaggcgaa	gggctggcat	gccaagcgtc	37920
ctagttaatg	cccagctgag	cctgggtgaa	gaaggatggg	ggtgtgggga	agacaccccc	37980
caccaaccgc	caaagacagg	cgcacaccag	ccagctctct	acttcccttt	ttatttccct	38040
taagacttgc	aagcagcagc	accagagagg	gaacctgccc	tcctggccct	ggaaggggccc	38100
gacccccaac	ccctaaccac	ggacacagct	ggcacctcag	gcccccttcc	ttctgaaagg	38160
agggctgtgt	ctctctcaca	ttcacacata	cacagacaca	tgcatgtgtg	cacactcatg	38220
gcacatggga	cctcaggggt	agcctgtttg	ccgatcccc	caagaggtag	caggaggcag	38280
accgctagaa	ggagataaga	ggcaccctgg	tctcctccaa	cccaaggagg	aagaaagctc	38340
aaccctcta	ggataggggac	tgtcttcagt	caatggagcg	ttgacttagg	gggcgttttt	38400
gaaggttttt	tttccctcct	tttgagctct	ttacaaaaat	agaacttctc	ttgggtattta	38460
taaatctacg	gccatggctc	tatgtgcag	ttacaggtag	aaaagccata	tggggcactc	38520
cttttggttg	ctcaggcctt	gattgcctgt	catccaggct	ccttggtctg	agaagtctat	38580
gcgtcacct	cagagccgct	aagcaccttc	agtgggcccc	tccattggc	ggcgactacc	38640
tgctggagcc	gggcacggta	atagaagagg	taggaaggca	acaggaatcc	caggagttag	38700
aatagcagga	ggcccagatt	cacctttagg	gcaaggagag	agaaacagag	tcaagtaggt	38760
agtcactctg	ccttagcctc	ccacagggag	gagaaggcgg	ccatttttct	ccaggctcctg	38820
agccagaata	aatacagcta	gtacttatta	tgtgtagtca	ttgttccacc	agtatctcac	38880
ttaagtgtca	gcaattctgc	aaagtggctg	agtagagact	tctcaggtag	aacaagtggc	38940
agggcctggt	gggtgcccac	acccatgggc	actcactagg	taggtatgag	gaaggcacag	39000
cactgtagga	gtctgggctg	gtcaggctgc	tcccgaatg	gggccttctg	ggctacccc	39060
tctgaccttt	ggagatgtta	accaatggga	tcccgttcag	ggtggcgaga	ggaggctctc	39120
agacacagtt	caagggaactg	ggatgcacag	cctggtggac	agaaggcttg	gaaggcccag	39180
gacacgcggg	ctctgactcg	gttcacatcc	cactctgcat	tactcactgt	gtgactttgg	39240
gcaataaatg	gcaattctta	ctgagtgcct	ccttctcagg	gctgttgtgg	cgaagatgta	39300

agttaaaaaa	aagtatgcat	catgcttagc	acatagttag	tgcttggtta	atagaagcag	39360
ttattttcatc	acaatttcttt	gggaggaggg	tttacgtgtg	gggtggcccca	cagggcagat	39420
gaaagatcag	cgtcagggag	gcagatgagt	tcaatgtaag	gaaaagactt	actaacagca	39480
gcagggctgc	ctcgtgcagg	agtgggtgcc	ctaccactga	gggtatctaa	gctaagaggg	39540
aaggggtcccc	tttcaggggt	gctggagaca	ggatcccaca	ctaggtagaa	ctggattgga	39600
ccaatgggtgc	ctgaacacag	gcccagaggt	caggactggc	cacttcacaa	agcacctgga	39660
gtttactaaa	aacagactcc	taggaggtca	ggcactgtgg	ctcacgcctg	taaccccagc	39720
actctgggag	gccaaggtga	gaagatcatt	tgaggccagg	agtttaagac	tagcctgtgc	39780
aacatggcaa	gaccctgttt	atctgtacaa	aatttttttt	taaaaaatta	gccaggtatg	39840
gtagccatca	cctgtgggtg	cagctactca	gaaggctggg	gccggaggat	cgcttgagcc	39900
caggaatcag	aggctgcagt	gagctgtgat	tttaccaccg	cactccagac	tgggcaacag	39960
aacaagacac	cttctctaca	aaaaaaaaaa	aacaataggg	ccgggcgcgg	tggctaaggc	40020
atgtaatccc	agcacttttg	gaggctgagg	agggcagatc	acgaggtcgg	gagatcgagg	40080
ccatcctggc	tagcacgggtg	aaaccccgtc	tctactaaaa	atccaaaaaa	aaaaaaaaaa	40140
ttagctgggc	gtggtggttg	gcgcctgtgg	tcccagctac	ttgagaggct	gaggcaggag	40200
aatggcatga	acccgggagg	cggagcttgc	agtgaaccga	gatcgacca	ctgcactcca	40260
gctgggcaa	cagaatgaga	ctccgtctca	aaaaataaaa	ataaaaaata	ataaataaat	40320
aaaataacaa	taaattaaaa	acaaaaacag	actcctacgg	tcaggctgag	atatcctgat	40380
tcaggggact	ggggaatctg	tatttttaac	actccgtgag	gggttctaaa	aggcagacaa	40440
cttggaacc	tgcagattag	agacctctga	ggtgcctctg	gctgagatga	gtgagggatg	40500
gcaccacata	caaggcccta	cccctgcccc	caggagagtg	gctcctgctc	ccccacacc	40560
aaccctcgct	ctcaccaga	aggcctctcc	tttcagggtg	cccaccatcc	ccatgaaaag	40620
tggctgctga	agcaaggcga	acacagcact	ggtgaggggac	tgcaggcctg	tcagcgtccc	40680
aaaaggggtt	ggatgggaac	ctgtcccaaa	aacgggagat	caaaggggtg	tgggggcctt	40740
tcagcccagg	caagaacttt	ttcttttctt	tcccaacatg	ggnnnnnnnn	nnnnnnnnnn	40800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40920
aagaaaaaaa	aatctttaaag	nncactccag	cttgggtgac	agagtgaac	cctgtctcaa	40980
tcaatgtttg	tatttcaagc	caagtgttat	tacaaaacag	tcaaaagtgt	ttaaaaaatt	41040
aaaagtttat	aaagtaaaaa	agctaagtaa	gctaggggtta	atttttttat	cgaacaaaga	41100
aaaatatctt	tgtataaaact	tagtgtagtc	taagtgtaca	ttgtttttat	tttattttat	41160
ttttattttt	ttgaaatgga	gtttcactct	gttggtccag	gctggagtgc	aatggcatga	41220
tcttggctca	cggcaagctc	tgtctcctgg	gttcaagcga	ttctcctgcc	tcagcctccc	41280
aagtagctgg	gattataggc	acccgccacc	atgcatggct	agtttctttg	catttttttt	41340
ttgaaatgga	attttgctct	ttgaccagg	ctggagtgc	atgggtgcaat	ctgggctaaa	41400
tgcaacctcc	acctcccagg	ttcaagagat	tctcctgcct	cagcctcctg	agtagctggg	41460
attacaggca	tgcaccacca	cactcggcta	atttttgtat	tttttagtaga	gacaggggtc	41520
tcaactaaag	agaacctagt	tggccaggct	ggtctagaat	tcttgacctc	aggtgatcca	41580
cccacctcgg	cctcccaaag	tgtctgggatt	gcaggcatga	gccaccatgc	ccagccagta	41640
tacagtgttt	ataaagcctc	cagtagtgta	cagcaatgtc	ctagaccttc	acattcactt	41700
actactcact	cactcactca	cccagagcaa	ctgccagctc	tgcaagctgc	atgcatgata	41760
agtgccctat	ataggtgaac	catttttttaa	tattttttat	tatattttta	ctgcaccttt	41820
tctatgatta	gctacacaaa	tgcttaccat	tgtgtttacaa	ctgcctacag	taatcagtac	41880
agtactatgt	atgggtttgt	agcctaggct	ataccatggt	gcctacgtgt	gtagtctct	41940
atactgtcta	gtttgtacac	tctatcatgt	ttgcataaag	ataaaatcac	ctaagtacac	42000
atctctctga	gtgtattcct	gttggttaagc	aacacatgta	taaacattta	caagaaatag	42060
ctcaaatttt	tttttctttt	gatacagggg	cttgcctttgt	cacccaggct	ggagtgcagt	42120
ggcgcaatct	cggcgcactg	cgacatctac	ctccccgggt	caatcgattc	tccggcctta	42180
gcctcctgag	tagttaggac	tacaggcacg	caccaccacg	cctggctaatt	ttttttgtat	42240
ttttattaag	agatgggggt	ttgccatggt	ggctaggctg	gtctcgaact	cctgacctca	42300
ggtgatctgc	ccgccttggt	ctcccaacat	gctgggatta	caggcatgag	ccaccatgcc	42360
cagccattac	gttttttttg	ttgttttaatt	tttttttttt	taagagacag	attctcactc	42420
tgtcatcaag	gctggagtgc	aatggcacaa	ccatagctca	ctgcagcctc	caactcctgg	42480
gctcaaggga	ccctcctgcc	tcagccttcc	cagtaactga	gactacaggt	gtgagccacc	42540
atgctcagct	aattattttt	tatcttttat	tttttgtaga	gggggggtct	ttctatgttg	42600
ctcaggtttg	tctcaaactc	ctgggctcaa	tcaattctcc	tgctttggcc	tcccaaaggg	42660
ctgggattac	aggtgtgagc	ctgaaaacct	tctagtctgg	aagtggaaga	taggccagg	42720
ccacttatgt	tttcaagtta	agcaagggtt	aggtcactta	tgaagcctga	ctagttttgt	42780
ttgcttaagg	gatctgcagg	cctgacctcg	gttttcatct	gttttaacag	tgtctatgtg	42840
tatgtgtgtg	tttatgtacg	tgcattgatg	ggggaaagct	cagaaatcaa	gtaagccaaa	42900
cacaaacatg	taattataag	cagggataaa	ttctatgatg	aagaagtatg	ggccacggga	42960
gagtacttgt	gccagtctgg	tgatcaggaa	caatgtcctt	tgggaagtga	catttgagcc	43020
atgccctgaa	gtacggttag	agttgggttag	gggtgaggga	gtaagaccga	gagctggggc	43080
ttcctgcaca	agctcagctg	ggcactgagg	accagtgga	ctctgctaca	gggcagtgag	43140

gagcagaaag	gctgaggaag	gctgggtgtg	gtggctcaca	cttgtaatcc	cagggccttg	43200
agaggctgat	gggggaaaat	cggtagagct	caggagtttg	agaccagcct	gagcaacata	43260
gcaagactcc	atccctgtaa	aaagctttta	aaaatttagct	gggtgtgggtg	gtatgcatct	43320
gcagctccag	ctactcaaga	ggctggggta	aggattgctt	gagcctagga	ggtggacgct	43380
gcagtgcgcg	acgattgtgc	cactgtactc	caacctagga	gacaaagcga	gatcctgtct	43440
caaaactgaa	tgaataggct	gtgtgcggtg	gctcactcct	gtaatcccag	cacttttgga	43500
ggctgaggtg	ggtggatcac	ctgtgattgg	gagtttgaga	ccagcctggc	caatatgggtg	43560
aaacccgata	caaaaattaa	ctgggcatgg	tggctcacat	ctgtaattcc	agctactcgg	43620
gaggctgagg	catgagaatg	tcttgaaccc	ggggggcaga	gggtgcagtg	agctgagatc	43680
gcaccactgc	actccagcct	gggagacagc	gagactccat	ctcaaaaaaa	aaataataat	43740
aataacaatt	aaaaaaaaat	taaaaggcca	gggagcactg	gcagcctgtc	caagggtttca	43800
ggtcacttta	gtaaaggag	aacaatggct	cctcccagga	cctctgggat	ctcagcattg	43860
atacgacagt	catggaaatg	ctagggccca	ggcagaccat	ctcagggaaa	acaagtggct	43920
ctgccctgcc	ttggccactt	cctggccctc	tgcattgccc	agggtctcag	caccaagctg	43980
ttctcagtga	gtagctctca	tttagtgcca	gggctctcgg	gcttacatcc	tacgatgacg	44040
atggaatgca	taaaagatgg	ggctgtgata	ggccagagct	aggggtttga	atctcatgag	44100
atgttcatgg	agccctggga	gggagctcag	tgcaagttca	tttctctttt	ttggttgaga	44160
tggggctcag	aggaggaagg	acttggtcaa	agacacacag	ggagtgtttc	agtgtgggac	44220
ggaggtttat	ggagaaagg	tgaccatcca	aggcttggac	aaagatcatg	acttcgacca	44280
gcaagcctca	actctgtaga	cttgggtggg	gccagccctc	cccaaacaca	cctgacaggt	44340
gtctgtggtc	ttggggacat	tgtcgctccc	cttctgctg	atgctctgct	gtccctctcc	44400
catgaagcgt	atctcttcgc	cgtcccccac	ccttgctgag	agaggatggg	ttctcttctg	44460
accaatactg	aagatcttta	gtaaagttct	cttttttttc	attttctgaa	agtccctctc	44520
ttgagaaatc	aggacaagtg	agtcagggcc	aggacaaaaa	acagtgtggg	acgagtgtgg	44580
tggctcacgc	ctgtaatccc	agcactttgt	gaggccaagg	tggcggatca	cttgaggtca	44640
tgagtttgag	actagcctgg	ccaacatggg	gaaacctcgt	ctctacaaaa	tacaaaaatt	44700
agccaggcgt	ggtggtgcat	gcctgtaatc	ccagctattc	gggaggctga	ggcaggagaa	44760
tcacatgaac	ccaggaggcg	gaggttgcat	cgagctgaaa	ttgggccaact	gcactctggc	44820
ctcttggcaa	cagagccaga	ctacctctca	aaacaaaaac	aaaaacaaac	gacaaacagt	44880
gtagactttg	tgtttttctc	aaaagcactg	tcaagccagt	gcccgcagca	gtgggcctag	44940
acacctccag	tcttgccctc	gggtcagttt	ccagcctccc	tggacacttc	ccccaggtat	45000
gtgtactttt	tgattgtcct	aaatccagag	tctgtggcct	gacctgggtt	gtcacagctc	45060
tcagtccttc	cccatcccga	atcccaggga	gcgcaggtg	tgtgcagaag	aggcacacca	45120
cactcaatac	atcttgcctc	ctcgctggac	ccaatccatt	ggcttgggtg	tgtaacagact	45180
gagcctcatt	atagccgttc	gttcctgttg	acctttccag	atcaatctgc	cagcttggct	45240
tctcgagttt	tcgcttgtca	gcatttctcc	aatccccatc	tgtactttgg	acctctttgt	45300
tgggtggcct	gctttatctg	aaattttcag	atltgacttc	aggtctctcc	tttgtccctt	45360
aatatggctt	aatggtggac	cctgtcaggg	gtagagaaaa	tattgaggag	ccctgacttt	45420
gaggtgcaca	agtttagagg	ttagacaagt	ccagccacaa	ccagcccaag	ctgcagtgtg	45480
gggaggcctg	tccagctgct	ccacggttga	gggtggagca	tacaggaagg	cttccttctt	45540
ctgcagcccc	aggtgttctg	gctgccctag	ctgcctggct	ttggtagaag	aaagaaaggc	45600
tctgtctctg	acttgtcaac	taatggcact	atgagattgc	acataattaa	cctgggtctg	45660
ctcttccaaa	agccttgggc	ctctgactgc	aacatggagt	ctgggtatca	ctccccatcc	45720
ctgcgccact	cacctgctct	ggcgctaggc	gtgtgcctaa	tcacttaatt	tctctgtgct	45780
gcctcttagg	tatcacttcc	cctgatccca	aatacttacc	aggtgtggga	tgacacctga	45840
ctagttactc	cttgaggta	tctgcttctc	accggggact	ccgaaaccaa	acgaaaggca	45900
aggccaagcc	cagcctaaag	gacgcttctc	acatgacttc	aggcttgcgg	gggtggagc	45960
gtgggggtgg	caatggagtt	ggggggggct	cagggagggg	atgtggaagt	gctttgtctt	46020
gcaaaactcta	gagaaccgtg	taaataggag	tgattattct	gtcccttccc	tttctttcca	46080
acaggaatca	gcatcccaca	gcccattgtc	agctatgaag	aatggaaact	gaggctccgg	46140
gagggttata	gggaggagcc	agcagggtct	tgagttcata	ttagtgcctt	ttcctccata	46200
ggcacatctg	tgttttcttt	tattttattt	tgaatttaat	tttttttttt	tttggcagag	46260
tcttgctctg	tcgcccaggc	tggagtgcag	tggcgcggtc	tcagttcact	gcaatctccg	46320
cctcctgggt	tcaagtgatt	ctcctgcctc	agcctcccga	gtagctggga	ttacaggtgt	46380
acaccaccac	acccagctga	tttttgcaat	tttagtagag	acagggtttc	acagtgtttg	46440
ccaggcttgt	cttgaaatcc	tgacctcaag	tgactctgta	gcctcggcct	cccaaagtgc	46500
tggtattata	ggtgtgagcc	actgcgctcg	gccacatctg	tgtttttaaa	gagaggaag	46560
gggataatgt	gcattttgtg	gaagcttggg	ccgttttgtg	ctaggactct	tatgatcttc	46620
ataagttttc	ccccaggag	gacactgttc	cacttaggga	gtcaggaccc	ccagtcctta	46680
caagattcag	cctctcaaaa	tggagacagc	agttccaggc	ctgggctggg	ttctgttcac	46740
actaggagag	ggcaagttag	tgggtgtttg	gatgtgggga	agtattatga	aaacagagat	46800
gttccaattc	ctagttagat	gaaaccatta	agctacttgg	catcttaaaa	ccaagagcgg	46860
ttcaagttct	gagattgtta	acacacctta	caacaccgcc	gccgttatta	ggaagaagct	46920
ctgtttgatg	acgtcccaca	ctgtgggtac	ctttatgaac	aggaatttgc	tttttcaaat	46980

cccagagaag	taagattaaa	gttggtgtgt	ctccatcctt	gaaaaatttg	gttttagggg	47040
gaattcaaga	atgactgacc	atacagaatg	gggagcaaac	ttgggaagaa	agaaggcaca	47100
gttcagagct	ctcccaatag	tcacccttga	actgcacccg	gaccatcagt	tatctctgtg	47160
ggtagagctc	aggaatctaa	aatccatttt	aaaattaaag	tatatcgggg	ctgggcgagg	47220
tggtcatg	ctgtaatccc	agcacttttg	gagggcagg	tgaggagatc	acgaggctag	47280
gagtttgaga	ccagcctggc	cacatggtga	aaccccgctc	ctactaacia	tacaaaaatt	47340
agccaggcat	ggtggcagac	acctgtatgc	ccagctattc	ggaaggctga	gtcagaagaa	47400
ttgcttgaac	ctgggaggca	gaggttgag	taagccaaga	ttgtgccact	gcaactccagc	47460
ctgggcaaca	gagggagact	ctgtctcaaa	aaaaaaaaaa	aaaaaattaa	agtatgtcat	47520
acatactgtt	acaggcacag	accttaagt	tacagcccaa	tgaaatttta	cacatctata	47580
cagctatata	actaccacct	atatcaagac	acattccagg	aactcagact	ccatcatacc	47640
cctcctcagc	agaggtaaca	gacccacacc	tctcctgtct	cggtggtaat	taaccactat	47700
tctaactttt	ctatcaatta	gttttgccca	ttcttgagct	tcacacagat	atacattgtc	47760
aggcatgatg	actcatgcct	gtaatctcag	cactttggga	ggccgagacg	ggagtatcac	47820
ttgagccag	gagttggaga	ctactctgga	caacatagtg	agacccccga	ctctacaaa	47880
aaaataaatt	agctggtcat	ggtggtgcgt	gcctgtatgc	ttagctattt	gagacgctga	47940
gagaggagaa	tctcttgagc	ctgggaggtt	gaggtgaag	tgagccgtga	ttgcaccact	48000
gcactgagc	ctaggtgaca	gagtgagatt	ctgcctcaaa	aaagaaaaaa	tatggccggg	48060
cgcggtggct	caagcctgta	atcccagcac	tttgggaggc	caagccgggc	ggatcacgag	48120
gtcaggagat	ggagaccatc	ctggctaaca	cggtgaaacc	ctgtctctac	taaaaaataca	48180
aaaaaagaaa	gaaaaaaaaa	ttagccaggc	atgtgtggcg	gctcttgtag	tccagttac	48240
ttgggaggct	gaggcaagag	aatggtgtga	acccgggagg	cagagcttgc	agtgagccga	48300
gatcgaccca	ttgcactcca	gcctgggcga	cagagtaaga	ctctgtctca	aaaaaaaaaa	48360
ggaaaaagaa	aaaatatata	tacattgtgt	actttttggc	atctggttta	ttttgctcaa	48420
tatcacatct	gcgaaattaa	tctacactgt	gtgtatgaaa	ggttggttct	ttttggtgtg	48480
atgcagtatt	ccgtcgtgtg	actacgggac	aatttgctta	tccgtattcc	tatcggtggg	48540
catttggtgt	gttaccaggt	tctggtcgtt	atgaataaag	ttgctatgga	tattctgtga	48600
cactacttct	ggtgagcgt	tgactcatt	tcgcttatgt	aaatatcttg	ggtggaatta	48660
cctgatcata	aggtaggtgt	gttggtttg	taatgtgctg	acttggttat	gctgaattcc	48720
cttttttgtg	tatttctggt	tagagcggaa	catgagggtg	tctcttcagg	gaatctggag	48780
ggtggaagg	aagcaggagt	cggtttctgt	ctcacacatg	ttgtgactga	actgctggta	48840
cacctggttg	gcatggagct	ggcttctcct	ttggcggtgc	ctactgttgg	ggcagggtg	48900
tatgtggtta	gctccatgca	atgaaccgg	gcttctgcaa	aatacattaa	caacgacaga	48960
gacaacaaaa	gctgatgtgg	atttaaagc	ttcagttcan	nnnnnnnnnn	nnnnnnnnnn	49020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49320
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	50820

gctaaacttt	ctccaggggg	taaagttcaa	gttgattagt	caattgatta	attaattcat	50880
tatgtaatgg	aaaaactcct	tctatgacct	gggcagagtt	ataggcagtg	aacaagacag	50940
acaaggtcct	tggtgtcatg	aagtttgctt	tctgaaggag	agagataata	aacaagaaac	51000
cagtaagaaa	gcaagattat	atcatttttg	taaatgttct	tgtggaaata	aatgtgatga	51060
tgtgtaacaa	aagtaccaa	taggagagtg	gggtgggtgg	gcttctttta	gaaagagttc	51120
tcggagaagg	cttatctgag	gaggtggcct	tttaaccagt	acaaatgctt	tagcttggcc	51180
agtggagctg	ggaccaggat	gacaagggtc	acttgtcatg	ccagtgaagt	tgagcttgta	51240
gacaagagcc	tgatcatgaa	agactttgca	gatgggtggt	atgggttttg	gttaattgct	51300
actatgtggg	aagactttga	atgggaagca	tggggacaat	ggcctgtgat	acatgttatc	51360
aaatatggtc	gcaggggcta	gtgaggtggc	agcagagata	gggagaagta	gacggactgg	51420
ggaaggtaga	agatggggca	ggggaggcaa	ttactgcaaa	gacatattcc	ttctaagctc	51480
actgagtgtt	catggtctct	gggagcagag	gttcctggag	gggaaagagg	ataatgtcac	51540
ttcctgagga	agcgggaaga	acccatctga	gacgtgggga	ctgtgctggg	tcgtttctaa	51600
ggggccttcc	agatctcaca	tgccaatcgt	cttgggtctat	gtcaattggt	ggggcatcca	51660
aatggggaac	tggtgtccag	gccgatttca	cagaacaacc	gccagtccta	tatctcccga	51720
gccattcaccc	cttgcaagtgg	cgtttagctct	ttcaccagct	tttatctgcc	ccgtggggat	51780
gttggccaag	cccagttaac	aagcagttga	tcagccccag	agatcaggtc	cctggagtct	51840
gtcacttttc	tgaggggtggg	gagagaatcc	tggagcagaa	catgtaacta	gaagggccac	51900
ctggcttcct	atgggtctgag	ggagagaatg	gtgggatctc	tggcctgaat	caaacctccc	51960
tttctcagtg	tccatcttac	ctctctgctg	taccttcggt	attttccagc	agctcctcag	52020
cccggttcctg	tgggaccttt	ctctgccaat	ccctacaccc	actgtaaat	tcaccgtggg	52080
agggagatgg	gccttgagggg	ctgtattagt	cttctattct	gcataacaaa	ttgcctcaaa	52140
tttagcagct	tcaaacact	catgtttatt	agctcatcgt	gagttcatca	gcagtgtggg	52200
cccagcatgg	ctagggttttc	tgctcagggt	ctcacaaggc	taaaatcaag	atgttgtctg	52260
ggctgtgtgc	tcatctggag	tttagggttc	tcttccaggc	tcacgtgggt	gtggcagaat	52320
tctgttccct	ggagttgcag	ggctgaggtc	ctgttttctt	gctgactgtc	agatgagggc	52380
tgctctcagg	tcctcgaggc	tgcccacatt	gcttgccacg	tgcgtggtct	tttccatcct	52440
tgaagccagt	gatggagaat	ttcccttggg	ttgaatcac	cacatgggtg	gactctctga	52500
cttcaggaag	agagccctgt	ctcttttatg	ggatcacctg	attagatcat	acccatagag	52560
ggcagttcct	tttccctaaa	gtcaactgtg	gcatgtaaca	tcacacaacc	acaggagtaa	52620
aatccatcat	atttacagtc	ccagggatta	tgcacagtgc	accaggggac	aactgaattc	52680
tgctgtcaca	aagggccaa	caggacttta	ttggtgaaga	acagtggaa	gtcattcttg	52740
gttcttccag	aaaaaaatca	ctcagtaaa	ttagaggttc	tcttgccctt	tgggaagtca	52800
tcaaagaatc	tcatggaggg	tttggacctt	caccctagaa	acatcacacc	atgttttcta	52860
taattgcagg	gttcatggtc	ccttgaagcc	tattcatagt	ttccagggtg	aaaagctctg	52920
ctgcagggtg	tggggagggg	tgcaggtgga	ggtgagggct	gaatagtgtg	agctgcatat	52980
ctggagctgt	ggtggttttt	ttagtcttta	agctgtcatg	tgttgggggt	tgggcattgg	53040
aggggcatcc	caagagctcc	ttggtattga	caccatctcc	aagggtgatct	ctgctctgcc	53100
tgggtgcacac	atgtttttct	cctgttgcaa	cagccactc	ttgtagaaga	gcagaccctc	53160
cagtaccagg	tctgaccctg	gacagcttgt	accaggagct	acagcacact	ccccacaag	53220
cctaaagtgt	ggatgagccc	cccagagaat	agatcagaaa	agattaaatg	cagagggtgat	53280
ctgcaggctc	ccctttggaa	gtgctgggat	ggagaggatt	gactgagctc	gtttagggaac	53340
ctccaagctc	tgtagtaact	ttagggttag	aaaggaggat	gcctaagatt	caggatcctg	53400
cagtgatgag	tcaacatttc	ttggggaagg	aggcagggtc	gaggattaaa	cggagatgat	53460
gggtatcggt	ctcttgctca	aaggcactgg	accccaaggc	ctccagctct	tcgctcccat	53520
ttgaaattca	agtccctgagc	acaccacagt	tgtgatgcag	ggaaagaatg	tgcttatcag	53580
agagcctggg	caagtgggcc	ccttggtgag	accgttcaac	ctcatttatg	tcattggcac	53640
caaaagtaga	catcagctc	ttgaaaagtt	gattaatgct	ggtcacactc	aaagaccctg	53700
ggtagcattc	atttactaag	caattactaa	ataccagttt	ctgtgctaaa	tgctgcatca	53760
gtcagggctc	ttaatggcag	gcagcagaaa	ctctccttgg	ctgatctaag	tagaaaaatc	53820
caggactgaa	aggaaacgga	gtagctcatg	aaattgcagg	aagggccgga	aaaccagaca	53880
tggagccaaa	gtcaggctgc	agaacaggtc	tagggaggat	cccactgctg	ctgagaccta	53940
gaccttggtg	ctggcaccca	ggatgttgta	gggctcagac	cctggatcaa	tgtatcctgc	54000
agtgcctctg	tgggtactgc	aactccagga	actcaatctt	gtcaacgcca	ccgccagaga	54060
gaggccttct	tggcctccat	ctttttgggtc	actagctcca	gattcaaaat	cttgaataga	54120
tgcttcttct	ctttgataga	gcccgatcat	atgctgttagc	tgcaaaggaa	gctgaaaatc	54180
tattaggaac	ttttgtcttc	aaaaatgaga	ggcctgtcct	ccaccaagat	ccataggaaa	54240
tggaaatcaa	gaaaccacag	gaagggtgga	ggtgactggg	cagctcacag	catgcatgct	54300
acatgtgaat	tatctcattc	atctctcaca	ctacccagtg	aggtaggtat	tgtcatccct	54360
acttcataaa	tgatgatatg	aggtacagaa	agtttaagga	acttgcccag	gacacgcac	54420
gcagctatta	agtgctagac	ccagtcaatt	tgagtctgac	ttggactgtc	tgactccaga	54480
agccaccctc	tcagacactg	ctgtatactt	ccagtgaatg	ttgatgaaat	tttcagggtt	54540
gctaagctgt	ggatttcaga	tcctggattg	tatgacctaa	aagagagact	tccttaggag	54600
tgagggctcc	tgaacagtc	actggtttcc	aagaatgggc	tcctctcat	caccttatga	54660

cagtaatcct	ctgtccaaca	gccaaagagg	tcctgtgggg	agggccttgca	gatgggagtg	54720
cgcagagccc	agctcaaagc	tcctgactag	gctcctgttg	agtattcctt	tgattcctgc	54780
ttctgtcttt	ttaaatcaat	ggagacaggg	gaggggtatc	tccatcctcg	gctcaagatg	54840
aaatgcatcg	ttcctcgttt	ttctcattcc	ttcccaatgt	gtgtactggt	aacttttagtt	54900
atgaaggaaa	ttacagtgtc	ctgtgcatat	accaaggctg	tccaacctcc	acacctttgc	54960
tcaagctggt	ccttctactt	gaaatgcctg	tttcttcccc	ttctaattgc	atctttccat	55020
ccaggtagga	atcagctcct	tggttcattg	agccttttct	gctctgtttt	actatgcatg	55080
gacttccttc	tgaattagca	gaggatgttt	cctagcttgg	tcttaaccct	tctccttttg	55140
tttgacctca	atttactcat	cttacaatt	agggtgtga	ctaattgaat	acaggatcta	55200
tgcttcactc	tgattttatc	tccacctgga	tagcatcatt	tttgaacac	aagcaggcat	55260
atgggagggg	agagaagttt	ggtgccagaa	agaactggat	ttgaattcta	accctgttgt	55320
ttacgtgagt	acgttactta	accattaatt	acttcaatgt	atatttatta	agtacctact	55380
atgtgccggg	cactgtacta	agcaccaagg	atacaatgg	gagtaaagag	atgcagcctt	55440
caccatcacg	aaggaagaca	gatgttaatc	cattaaccaa	gtaatctcac	aagaaaagta	55500
aaatgactaa	ctgataagga	caagccccctg	gagctacaag	aggggtgtata	cagggcatcg	55560
atccaataag	ggcagtgttg	cggggagatc	aggagccaca	cagagcctgg	gttgtctcac	55620
ttggaaaatg	gggtatcaac	cacctacctc	actaggtttt	taaaatcagg	ttaaatgagg	55680
taatacttgc	catgaacagt	attttgttga	ttgatgattg	attgaaacgg	agtctcactc	55740
tctcgcccaa	gctggagtgc	agtgggtgcaa	tctcagctca	ctgcaacctc	tacttctctg	55800
gttcaagtga	ttctcctgcc	tcagactccc	aaagtactgg	gattacagge	agccacctct	55860
atgcctgact	aatttttgta	tttttagtag	agacaaggct	ttgcaatggt	gaccaggctg	55920
gtctcaacct	cctgacctca	aaagatccac	ccacctcagc	ctcccaaagt	gctgggatca	55980
caggcatgag	ccactgcata	cagccacttg	ccatgcatgg	catttaaaaa	tgttcagtaa	56040
atgttaccat	aatgaaggct	ggtagggttg	ccaactgagt	ggtctgattc	agaaggaaaag	56100
aagttagaca	tacgtgaaca	tttctgttac	ttgaagatcc	tcaggacagt	gactcctaga	56160
cccactcttc	atcacagtca	gctgggaagc	ttttaaaaa	atgcagacat	ctgaccttca	56220
cgctagacct	attagccaag	cagaagtctc	tgggcagggc	atctgcata	ttttaaaaaat	56280
ctttaataag	gcagcctcaa	aattacagat	tcagcacgca	tttaccataa	ccactgaaga	56340
aatgcaaagt	tataaaaaaga	agataaaaca	caatctgtct	cctgctttct	tccctctcct	56400
cccctgcttc	tgagggcaac	aaggctcaact	atttgggtgtg	attcctttta	gcattccctc	56460
catcaatggt	cacataagga	tgctcacaga	taagcaccta	tgcggggggt	ttttttttcc	56520
ttgtaaaaat	atcacataac	taataacttt	cctcagtatc	ttgccttttt	tcacttcatg	56580
tcacagaaac	atctcttcag	gtttatagat	acagggtccag	ctcttctttt	catagccata	56640
taacattctg	tagaatagag	aggacacatt	ttactcagtg	tcggattgat	ggatatcaat	56700
attgttttca	tttctacaaa	tagtcaagga	ataacataac	tctgtaaaag	ttttattact	56760
tataggcgca	tttatgccta	aaggatagtc	tcaaaagagt	gaaactgatc	aaatgtgcat	56820
ttttttattt	taataggtat	ggacagattt	gttctcaaaa	tgtttggtgg	agttcaaaac	56880
accagtaaaa	caggggagat	atgtattttg	gaaaagcacc	caaggcgatt	ctgaagtgtg	56940
gcccaggata	agaaccattg	cccagagctg	ttccagatgg	ccctggggtt	cctgaagtgg	57000
gtatcgggag	agaaatcttc	actgaatgaa	tgagtgggct	ccccagggaa	gtgatgaaat	57060
ggtccttatc	agccttgcta	tctcctctcg	acagaggcaa	actctctctc	cctgggggaa	57120
gttccctcaa	ggcctctata	taagaagtct	ttgtgagagg	aagcaaaaga	ggacctgggc	57180
tttgggaaga	tctaaagacc	caggaaggtc	tctgggtggg	tgagtgcctt	ctctgctgtg	57240
gtggagctgg	tgacagttta	ttctcccagg	aggtccctgg	ctgtggctga	cagtttctgg	57300
agggctggca	ggcgtctacc	tgtggctttc	aggttatgag	gatgtcagca	ggggcagcct	57360
tcactcctcg	ccttgacat	tccttctcg	ggatgtgaaa	gtgctccttg	gctggggaaa	57420
ggagatggtg	gagacatgga	ggagggtgtg	ggtggcttct	tgaactctga	ggaggggaca	57480
taccttctaa	gtcctatgtg	ttcctaggaa	agccaataat	cattgcttct	cccgcctttt	57540
ttatgtcata	gactctgagg	gacccattaa	gtacaaacaa	ataagcgtaa	tagtcccttc	57600
tttacttccg	ggcctgaagg	aaagccagcc	tcagccaccc	ctcagggttt	gctgcgttct	57660
gtttagaaa	aggtccttgc	gtcctggatc	ctggagcatc	aggagctggg	cttggcatga	57720
gcttttctgg	cccactcctga	tttctattca	ggccttcttt	ttctccacct	cactcccacg	57780
gtccccta	ggtgtgattg	tgatgtgtgt	gcattgtgtg	ctgtgtgtgt	caatgacaaa	57840
ctgtgttctc	cggtgcagga	taaagccaag	atgaaactcc	ccttacttct	ggctcttcta	57900
tttggggcag	tttctgctct	tcacttaagt	aagtgttttt	tgccttcagt	ctttctttct	57960
ctgttttttc	cctttctatg	gtagatgggg	tcagagttac	acacccaccc	ccttctttga	58020
tcgtcttcta	tttctgaatt	tctgtgtgct	taaagggatg	gggactctat	ggccaggagt	58080
tgaagagatt	tctcaaggcg	tctgttatgt	ctgtgttctt	ggttctactg	tgacattccc	58140
aattttgtcc	tttctccatt	atgcttactt	tgagcttact	gagtgccttc	tctcctttta	58200
ctctcttagc	atcgccatga	agtaggtggt	attgtatacc	catttcacag	aaatacagct	58260
ggtggatgat	ggaaccagta	cccaagccca	tgactgcccg	actctaagtc	catgctctta	58320
accaccttga	ccttgtcagg	cagcttgggt	ccccctcata	gagactgggt	tccaggttcc	58380
ccttcccagg	cagagttgag	cactctgatg	cccagggcaa	ggtgtgagct	gtctgtgggt	58440
ctggggagga	acaaggggag	atgtgaagga	aggacactta	gctatcctcc	ctgccagggt	58500

ctgagacttc	cacctttgag	acccctttgg	gtgctaagac	gctgcctgag	gatgaggaga	58560
caccagagca	ggagatggag	gagacccctt	gcagggagct	ggaggaagag	gaggagtggg	58620
gctctggaag	tgaagatgcc	tccaagaaag	atggggctgt	tgagtctatc	tcagtgccag	58680
ataitggtga	caaaaacctt	acgtgtcctg	aggaagagga	cacagtaaaa	gtgggtgggca	58740
tccttggtg	ccagacctgc	cgctacctcc	tggtgagaag	tcttcagacg	tttagtcaag	58800
cttggttgag	tgccctatgg	ctgaggctga	ggtgggagca	tggaacgggt	gtgggatatg	58860
ccccagcat	tgctatcact	ggctcttttt	cccattgagg	gccctggggg	tgtcagtaga	58920
acctgagcct	cagagagggtg	ttggggtaag	aggggagggc	cacctacaaa	cagaagtgtc	58980
attttggctt	ccaaccttca	aatggttggtg	gcaggggagg	gaggggaatga	attgtgggga	59040
ctcaagaccc	atgtgaattc	atgtaggaag	gatgtccat	tctttgtctt	ttatcctgcc	59100
ctgtagttta	cttgccggag	gtgctacagg	ggcaacctgg	tttccatcca	caacttcaat	59160
attaattatc	gaatccagtg	ttctgtcagc	gcgctcaacc	aggggtcaagt	ctggattgga	59220
ggcaggatca	caggctcggt	aagagaagtg	tgaacactaa	atgggggtgca	cctgctgatc	59280
tcagccagca	ctcagcttgc	atcagatttg	tctgttttcc	tccgtataaa	tctccagaag	59340
aaccagggat	agatggacac	ccacagacaa	cctgagggtg	gctgcctggg	cattcaggga	59400
agagctaagg	atttagaatc	aggagggttg	ggtccaagtt	cctttccatc	tctcactatc	59460
tatgtaactt	aagttagctg	ggcatggtgg	tgcatgtctg	taatcctagc	tacttgggag	59520
gctgaggcag	gagagtcact	ggaacctggg	agacagaggt	tgcggtgagc	cgagatggag	59580
ccattgcact	ccagcctggg	caacaagagc	gaaactccgc	ctcaaaaata	aataaaaata	59640
taaaataaat	aaaaaaaaaa	ttaaaacaag	acctagattt	tgtttctcca	tctctaggat	59700
gagttggcaa	cccttggtct	accttttggt	agggtggaa	ggacaagcct	gtcactggga	59760
tgcatagaat	ctgatggtga	taattgccgt	ggatcagcat	ttcagatgac	taggacagtt	59820
cccatcatgg	tccagcaggg	aagggcccat	tgcccggtgg	gcagcagaaa	gagctggcag	59880
atacggggcc	aggtctgctt	ctctgccttc	cctctgcccc	atcccttctt	cccctcttgc	59940
tttctccagg	gtcgctgcag	acgctttcag	tggtgtgacg	gcagccgctg	gaactttgca	60000
tactgggctg	ctcaccagcc	ctgggtcccg	ggtgtcact	gcgtggccct	gtgtaccgga	60060
ggtgaggtgg	ggctggggat	gaacgatgga	aaggtctggg	agatgggaag	tgccccaagg	60120
aggagatgct	acaaagagcc	tgaccttttg	tgggagaggc	ttcctgggtc	ttttatatac	60180
tctgactcca	cagcagtgtg	tgggtgggaa	aagaggccct	cctgtgggtt	gagttgggat	60240
ggacaagagg	ctgaaaagtc	ctttctgttc	tgctttcaca	ggaggccact	ggcgtcgagc	60300
ccactgcctc	agaagacttc	ctttcatctg	tctctactga	gctgggtcca	gccagcagtt	60360
cagagctgcc	ctctcctggg	cagctgcctc	ccctcctctg	cttgccatcc	ctccctccac	60420
ctccctgcaa	taaaatgggt	tttactgaaa	tggttttatt	ttctcctctg	atcgcggate	60480
cactctgctt	agccctcatt	gaaacttctt	ccttatcatc	tctccccaca	ccacaacttt	60540
catagaagtg	tcagaagcta	ctactccttg	aggaggagga	tggaggggtg	agttgggtct	60600
atggagcctt	ttggagatgg	aggaatgggc	tcagctagtt	ctcttcatag	aacacctgat	60660
tactgggcac	ctgcatagtg	ctgccaggac	ctttcaaggt	tgtaggtaga	ctcccaatgg	60720
cccagtttgc	atctctgtaa	ccaaaggcct	tttctctctc	tctctccaac	cccagaactg	60780
tggttggttt	tatatgtaa	gaagttaaca	tgtccctggg	aacagtccac	aacattcagg	60840
aatgaatgta	taagtaccgc	aatccccggc	ccctcaagtg	gaataaatct	aacatgtatt	60900
gggcaccatt	tcccagtggc	ctgctgtggt	agttggcctt	attccatgca	tttttatggg	60960
ctgccttccc	ttcctcaact	gcattctctg	ctccttccca	ctctctgcaa	ctcccaaata	61020
aacacttgta	cgcaactccc	tctctcagga	tctccttctg	gggaaacctg	atataagaca	61080
gcttgccatg	cgtcagaactc	tgaatgaggc	ctgggaatac	aagacatagt	cctctggcac	61140
ttgggatata	tggttatttg	taacataggc	acaaaaacat	ctactagttg	ttatcgctta	61200
ttgagcacc	acaacatacc	ccctgctgtg	gcaggcacct	tgccatagtg	acctcatgtg	61260
atcaataatt	atgagcccta	ttttacagaa	ccaggctcag	agaagttagg	atctgtcaaa	61320
agacttgccc	aagactgaac	ctctaaatgc	aactcatatt	gaaattcaac	tctgtctcaa	61380
agcatgttac	tttaaccctt	gtgcttttac	agctggctac	tctcccctta	tggtcacacg	61440
gggatgaagc	acggggggag	gaaagccaga	ctgtctcact	cttgggttca	tcttgggaca	61500
caggacacca	gcccagctgg	aggtgaggga	gctttaatca	gaggggaggg	aggaaggcat	61560
tctcaacccc	ttctgtacta	gggaggtcag	cagaagaaaa	taattcaatg	ttctaagacc	61620
atttttttct	ccagcattcc	tccaattcat	agatcttcat	atgggattag	gggctcagag	61680
aggggtgaaa	caagaactct	attttttttg	agtgtggtat	agagaaggga	tgctacttct	61740
ctaaggtcac	atagtaagtt	gagaaagaga	gagaaatcaa	actcaggttc	atttcaacta	61800
ttgttccaca	agaatctgtt	gattttcaaag	atgggtgact	atgggttcat	ccctgtggtg	61860
agtgtgtgta	ggatgcagct	gaggtggaac	tttcaactct	tgccctcttg	gactttatat	61920
tctggtgtgg	aaaggcattg	cttcccttat	ttcaatatta	acaacaaagg	gtaataatat	61980
ttcccattta	ttaagcattt	actaggtgtc	aggtactgtg	ctaaatgtta	ggtgaacttt	62040
gtcttgttcc	tcataaatct	ctgccgctgt	gggtgtgtac	tttgacagaa	gtttgacttc	62100
cagtcacag	agatcttctt	tgggggagta	atatcaagaa	ggggcacgaa	ggaagctgca	62160
gggctcctag	tcccatcctg	tatctcgacc	taggcattgtt	tacattggtg	cattcactgt	62220
gaagtttccc	tgagcagtc	actctatagt	gtgctttata	ggagcacatt	gtacatccat	62280
tgaaaaattt	ttcttggccg	ggcacgggtg	ctcatgtctg	taatcccagc	actttgggag	62340

gccgagacag	gcggatcacc	tgaggctcggg	agtttgagac	ctgcctgacc	aacatggaga	62400
aaccccgctct	ctactaaaaa	tacaaaaaaa	ttagccgggt	gtggtggcac	atgcctgtaa	62460
tcccagctac	tcaggagggtt	gaggctggag	aatcgcttga	acctgggagg	cgaagggtgc	62520
agtggagccga	gatcgtgccca	ttgcaactcca	gcctggggcaa	caagagcgaa	actccgtctc	62580
aaaagaaaga	aagagatttt	ttctttttct	taaaaagtaa	aaatcatgaa	ataaggggac	62640
tgggctaata	ttccaaaata	tgggtttgtg	tgtgaatttt	cctctccagt	aagatactaa	62700
ctaagctctg	tgaaactgtt	tatctatggt	tctttatcat	tgaatccttg	gagttcctta	62760
cactgtgcag	agcacagagt	aggggctcaa	tcaacagtgc	actcattgct	ttttcataga	62820
caagggccac	cctcactcaa	ctcatgtgcc	aggcatagtt	ctgagagcct	tgcttaagct	62880
gatnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnggtcact	aatttggtat	63000
agagtattgt	tcattgattt	catttttatt	tgtttctgat	ttttaagat	tgttttactt	63060
gttttcttcc	tattattatt	ttattttatt	tgtaaaaacat	ttacatatca	gacatttaca	63120
ttttcccaa	ggtaaaactg	tgaaacaaga	tataattcaa	gaagtttact	ttccctctct	63180
gtttcttgta	ccctttttcc	tcttcttttag	gtaaccattt	ttatTTTTTT	aaatataaac	63240
attgtgtagg	tgtatataca	tgtattagtc	tgttttctatg	ctgctgataa	agacctatct	63300
gagactggga	agaaaaagag	gtttaattgg	acttacagtt	ccacatggct	ggcaaggcct	63360
cagaatcatg	gcaggagggtg	aaaggcactt	cttacacggt	ggtggcaaga	gaaaaatgag	63420
gaagaatcaa	aagtggaaac	ccctgataaa	cccatcagat	ctcgtgagat	ttattcacta	63480
tcacaagaat	agcgtgggaa	agactggccc	ccatgattca	gttaccctcc	cccactgggt	63540
cccaccaca	atacgtggga	attctgggag	atataattca	acgtgagatt	tgggtgggga	63600
cacagccaaa	ccatatcaat	acatttccct	ctctttttaga	taaaaggtag	tactactgtat	63660
acactattct	gcagagtttt	tttttttttt	gatgtaaactc	tatcctgagg	gtgctctgta	63720
gcaggggacct	ctcatgcctt	ttaaccactg	cctgggtctc	cattacatgg	ctgcagcata	63780
gttgccacag	cattcctgta	ctgatgacta	tttggaattgt	ttccagtctt	ttgctattac	63840
cagtagtggt	acaaagagga	tctggctaca	tgttcagggt	ggggaggggc	agatgtgtag	63900
cctgtcagga	gggtattgca	gtaatccatg	actgagttaa	tggtagttta	aagctaggat	63960
gagtcagtg	ggttggagag	aagtgggcac	atttgaatga	tatgtaggag	gtgaatgac	64020
agcattattg	atgagtttga	ggtggggcat	gtggggaaag	gattcgagga	tgactcccag	64080
gtttctgttg	ggacagtggga	tggatagtgg	ctcctcccct	ttttccaatc	ttccttggcc	64140
cttcgctgac	ttctgttggg	ttggcctaca	gagagcttct	ttttcctctc	tgttcgccca	64200
ggttctctca	ctttggcggg	ggcctctctg	tcgacgggtg	cttcgctggc	cttgacatcc	64260
ctgctgtgcc	tgggcttcgc	cctctgtgct	tcagttccca	tcctccctct	ccagtaacct	64320
accttcatcc	tgcaagtgat	cagccgctcc	ttcctctatg	ggagcaacgc	ggccttctct	64380
acccttgctg	aagtggcctt	ggggcgggct	ctgtggagac	ggacacactg	gggcaaagag	64440
aagctggagg	taaagaaatt	gggaggcaag	gcggggcctg	gaggcagtca	ggtgcgggag	64500
actgggtttg	ggggcagggtg	tggagggggt	gagaccagag	gtggtgggaa	ggatagaaca	64560
ttcatgcact	tgagccttta	catctgcggt	gccctctccc	tctgttttct	acctggtgaa	64620
ctcgtattca	tcctctgagg	cccacttctg	tttcagttct	ccagggaaga	aatggaaaag	64680
tgtcttccct	tctttgtgcc	cttagtactc	tagtcttact	tcctttgcta	gtgcgtgcat	64740
tgtctggcat	gccatccatt	tacatgcctg	tcttttcttt	cctggtgcag	cctgcatgag	64800
ggtcctgtct	gtttttccag	ggccccgcct	gtgccttctt	ctgggttctg	tgggtcaaat	64860
gtctgagcag	agctgaagag	ggaaaaggcca	gtcaggtgtg	gttgaggggc	aggcctaggga	64920
caggggagct	ggggacaagc	ggccgacagc	ccccagaggc	caggcttctg	cttgagggga	64980
gggtccctga	agctcactgg	aacccctctg	gtttctctcc	ccagtttccc	ttcagagcac	65040
tttggaagc	tctttgggct	ggtgatggcc	ttgtcggtctg	tggtgtctct	gtccagttc	65100
cccatcttca	ccctcatcaa	aggctccctt	cagaatgacc	cattttacgt	gagttactggg	65160
aggatgggga	tccttggcag	gaggcctggg	ccttaggcct	tggctgcccc	aaatctggct	65220
gtgatggcct	gggtatgtag	catggtgagc	cttcccaaag	ggtctgtggt	attcaagtat	65280
ttggggcaaa	agtatttgtg	tgtgtgggga	aacagacatt	ctggagtagg	gtggggaatt	65340
ctcacgaaac	ttcaagcaaa	atcctgagac	ctcaaagggtg	tttcctgctt	gtggtgagtg	65400
caggcccaac	ctggcctctc	ccctaggccc	acacagggtt	tccacagttg	gccccaggga	65460
caggacctct	gtgctttcac	ctctgtgtct	ttacacctgg	agggatgctc	tgaggtcctg	65520
ctctaggagg	tggctgtgag	tctcctgtcc	tttgacagaaa	ctgaggctca	aagggttac	65580
ttacgtgttc	agaggcacca	gctaaggagc	aaaagtcaac	tttgaattct	gtgttttgac	65640
tactgcacag	ctctattttg	ctcatttttt	atttttaaaag	cagcaaatct	tagaataaggga	65700
gtttaaatcc	atcacttgga	gaaaagaaaag	actaaatggt	ttttgttttt	gttttgagga	65760
cacgatctct	ctttgtcacc	caggctggag	tgcatgtggca	caatctcggc	tcactgcagc	65820
ctcgatcttc	tggactcaag	cgatcctctc	atctcagcct	cctgagtagc	tgacactaca	65880
ggcatgtgcc	accatgccaa	gcttattttta	ttttatttttt	ttgatagaca	ctgggggtttc	65940
gctatgttgc	ctgggctggt	tttgaattcc	tggcctcaag	cgatccaccc	gtctctgcct	66000
tccaaaatgc	tgtgattaca	ggcgtgaacc	actgtgcatg	gccaaaagag	taaacttgaa	66060
atctgaggcg	aatgacttga	ttgtgacatc	aggtgacctc	gtaatcagct	gtgtattcta	66120
gctggtgcct	ctaccagctt	cccatgtgac	cttgaacatg	tcattgaatg	ctcgctaggc	66180

ctctgttttct	ttatctgtga	aatgggcttg	atattcctcc	tctaccccaa	ccgatagtgc	66240
agaatgaaaa	gtaactgaaa	gtccttcctc	cagggcacca	tagtgtctgg	gtgaaaagta	66300
gaatataaac	tcggtagact	tctggtccct	tcattgggtca	tggaaatggac	cagtgccttg	66360
ttcattgagc	aacagttctg	ttgttcagaa	ttcctggatt	tcacctcact	tctgctctcc	66420
ctgcaggtga	atgtgatgtt	catgcttgcc	attcttctga	cattcttcca	cccctttctg	66480
gtatatcggg	aatgccgtac	ttggaaagaa	agtcctctcg	caattgcata	gttcagaagc	66540
cctcactttt	cagccccgag	gatgggtttg	ttcatcttcc	accacctttg	aggacctcgt	66600
gtcccaaaaag	actttgccta	tcccagcaaa	acacacacac	acacacacac	acacacaaaa	66660
taaagacaca	caaggacgtc	tgcgcagcaa	gaaaagaatc	tcagttgcga	agcagattga	66720
tatcacacag	actcaaagca	aaggcatgtg	gaacttcttt	atttcaaaac	agaagtgtct	66780
ccttgcaactt	agccttggtg	gaccttgac	tccagggggg	atgacctggg	ggaggaagtg	66840
tgtcaactat	ttctttaggt	ctggttggtc	cogaagccta	tatgtgcctg	gatcctctgc	66900
cacgggttaa	attttcaggt	gaagagttag	gttgtcatgg	cctcagctat	gcttcctggc	66960
tctccctcaa	gagtgcagcc	ttggctagag	aactcacagc	tctgggaaaa	agaggagcag	67020
acagggttcc	ctgggcccag	tctcagccca	gccactgatg	ctggatgacc	ttggcctgac	67080
cctggtctgg	tctcagaatc	acttttccca	tctgtaaaaa	tgagatgaat	tttgggtgtg	67140
aaagttcttc	ctggagcaga	tgtcctagaa	ggttttagga	atagtgcacg	agtcaggcca	67200
ccccaaaggg	catgggagcc	agctgacctg	cttgaccgaa	ggatttctga	cagactatct	67260
ttggggatgt	tttcaagaag	ggatataagt	tatttacttt	gggcatttaa	aagaaaattt	67320
ctctcgggaa	taattttata	gaaaaataaa	gcttctgtgt	ctaaggcaac	tactgtttcc	67380
atctctctag	gctttgggcc	ggggctgtgt	gtgtgtgtgt	gtgtgtgttt	gtgtgtatgt	67440
gtatgtttct	gaggaggccc	tacctgggca	tgagagggta	gggaatctgg	ctacacatct	67500
agtgtggcag	ctggaccag	aggtggggca	ggaacctga	ctatgattca	ccccgctggt	67560
cctgggatgt	gggccagag	acttctctcc	ccaggaaacc	ctctgcttcc	tcttctctct	67620
cacatcctta	actaacttta	gcagaaccct	actcctcact	acacaccccc	agctagaagc	67680
gctggatgga	atcagaaatt	cctagtttga	gtttcaattc	tgccccctcag	cagctgggca	67740
agccccctaa	ccactctgag	tcactagttc	cccacctgca	aagtgcagtt	aatcatttct	67800
atctctgatg	gcgattgtga	gaatgtaaag	tcattgcaac	tgccctagcac	atggtaggag	67860
cacatgaggg	tttgctcctg	tgtttactca	tgaccttggt	ggaggacggg	ggcaaagagg	67920
gagaagttga	gggtgcagga	ggagagatgg	caggtgggtg	ggatgggaga	atctggggca	67980
cacctgctgt	ctcattccca	ccttgctagg	agagggacta	ggaaaagaaca	gtgggaggca	68040
gggggatggg	ggtggaaggc	aggggggtgg	aggcaggttc	atccatccat	tcattcaaca	68100
aatgtttatt	gagcacctgc	cacgtgtcag	geectgtcct	gggtgctggg	gctataaaga	68160
tgcagaaggg	tctgaaaccc	agctcttctc	tcttctctgt	gatgtcgggg	tgtaatttcc	68220
agggggccagg	agcctgggtc	tgagggcgga	caccaaagtt	ctagtgggtg	ctattagcag	68280
cgtttaaatc	taatggatgg	atttggtctt	gttaccctgc	tcaaaagctt	tcagcagctc	68340
cccactgtcc	acaggacaaa	aatccagatg	ctagcctggc	attcaaggct	gtcactagt	68400
tgatctcaac	ctctcccctt	cctctctttac	ctcctaccac	cagcggggca	gagcccacc	68460
ctgtggacca	agattcccag	tctctgggtc	tgtgtgtgca	ccagttctct	tgctgggtg	68520
gctcacccctg	cctcagcttg	tgaaatccat	ctggctctgt	gggatcctgc	tcaaaatgtc	68580
atcttctcca	aaaatcatta	ctcaggcttt	ccagcatgtc	tgagtccctg	gcacttggtc	68640
acacccttcc	tggtgactgg	catttgcttc	cacatcatga	ccctcccacc	ccttgctacc	68700
gcagcatact	ccaggaggca	aggtctgttc	tcgctggct	ctaattaatc	tgtgcttacc	68760
atccacatgg	taccagctaa	ttcttggtga	atgaatgatc	gttgaatgag	tggtattctg	68820
ttttggcctc	agaaccaatt	agaaggagcc	agaaaaacac	atgggggtgg	gggaggtgca	68880
gtgtggtgca	gtggaaaaaa	acccttctgg	aaatctcagc	tctgtcactt	actttgtcag	68940
ctctgtgact	ttggatggac	cacttctttg	tcagtatggt	gggagaaata	gacatgcctc	69000
tctgggctgt	tgtaaggatt	acaaattagg	tcagatgctt	ggcatgtggt	gggttgaaaca	69060
gatcacagct	agcattacag	atgatataat	aaagccaaaa	aaagatgcct	aatgtccacc	69120
agttggtgaa	cggacaaaag	aaatgtacca	tatttggtat	attatttggc	aatcaaaaaa	69180
agtactgaca	cctgctacaa	cacggatgaa	tcttgaaaac	attagactaa	gtgaaagaag	69240
ccagacacaa	gaaactgcta	atgattccat	ttaaatatga	aatatcgggc	caggggtgcag	69300
tggtcatgac	ctgtaatccc	agcactttgg	gatgccaaag	tgggcagatc	acttgaggcc	69360
aggagttcgt	gaccagcctg	gccaacatgg	cgaacccccg	tctctactaa	aaattagccg	69420
agtgtagtgg	catgcacctg	taatcccagc	tacttggttg	gctgaggcac	agaatttggt	69480
tgagcctggc	aggtggaggt	tgacgtgagc	caagatcgtg	ccactgcact	ccagcctgga	69540
tgacacagtg	aggttccgtc	tcaaaaaaaa	aaaaaaaaaa	ggaaaaagaa	aaaaagaaat	69600
ttccagaata	ggccaatctg	tagaggcaga	aagtagattc	atgattgggt	aggcctgggt	69660
gtggaggcca	tggttagtga	tggtctaatg	ggaagggttg	tcttttgggg	tgatgaaaat	69720
gggtggactt	atggtatggt	aattatacct	caataaaaact	gttattttaa	ggaagaaaag	69780
atgcctggat	tccccaggaa	gtgtacagta	gacttctgtg	agaatcagaa	atgatttctg	69840
gggaagatgg	gcgagaggag	agtaagtggg	agaagtgacc	acgtgcgcaa	ctctcatcgt	69900
tctgccctga	gagccttcc	cctgcaactt	tatttattta	tttattttga	aacaggttct	69960
cactctgtta	ccctggctgg	agtgcagtgg	tgtgatctca	gctcactgca	gcctcgacct	70020

gccaggctca	agcaatcctc	ctgtttgagc	tectgagtag	ctgggactac	aggcgcatgc	70080
caccacatct	ggctaattctt	ttattttat	atattttat	ttatagagat	tggggagtct	70140
cactctgttg	ctcaggctgg	tgtcaaacgc	ctggactcaa	gtgatcctcc	caccttggcc	70200
tcccaaagtg	ttgggattat	gggtgtgagc	actgtacct	ggcacctcct	gcaacttctt	70260
cctcaagtgg	aaccaatgag	gaagcaagca	actcagagct	ttcacaagtt	ttgatttcaa	70320
tcagcaacgg	gcttccaatg	caacccttct	ctcctgtaac	cagcctcagt	agagagggaac	70380
tggaggtgaa	ttggccccc	tcacaccccc	acagtgccaa	gctgggccc	tccatcaggg	70440
ggagaacaca	tgccgtgtaa	gggacagcca	acagcataaa	ataggaattg	tgtgatgatc	70500
ccttttaagc	ctattcagcc	cagggaagtg	catatgatca	gccccatttc	atagatgaag	70560
aaagtccagg	tcacccatta	gcacattgtg	gggctgggtat	ttaaaaccagg	tctgtctggc	70620
tcccaaggtc	acattcattt	agacattacc	tttactttac	atctcttctt	cttttcttct	70680
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cttcttcttc	ttcttcttct	70740
tcttcttctt	cttcttcttc	ttcttcttct	tccttcttct	cctcttcttc	ctcttcttcc	70800
tcttcttctt	cttcttcttc	ttcttcttct	cttcttcttc	ttcttcttct	tcttcttctt	70860
cttcttcttc	ttcttcttct	tcttcttctt	tcttcttctt	cttttttttt	tagggtgggg	70920
tcttcttcta	ttggccagg	tgaatgcagc	atcatcatac	ctaaatgcag	ccttgaactc	70980
ctggccttaa	gcaatccccc	tgccctggcc	tccaaaagt	ccaagatttc	aggcatgagc	71040
caccatgccc	agcctgcatt	tattctcttg	taagaaagat	atcatttaaa	acagacgaga	71100
aaataaagag	ggacatgaaa	aagacgcatt	accattaatt	ggaccactca	gagataatca	71160
tgggttaacat	gttgggtatg	tccttccctg	catttgactg	gatgtatgtg	ataatttaaa	71220
tgatctcata	agcttttctt	tatgtaatca	aatagtagcc	aaaaacatga	ttttaaatgg	71280
ctgctcacaa	ccccatctcg	tggttctgcc	acgccttggt	tatccccatc	caccccttac	71340
tccctttccc	cttccctgcc	tgtgtggggg	tcctagatga	cggtgagcca	gagggcagcc	71400
ttggtcagca	gattggagag	tgcaataaat	aaaaacactc	agaaggcgag	ctgttgtcaa	71460
gtgggcttat	cacaaaagag	caccttggga	tattccagag	aatgacctca	taccgcgtaa	71520
tcactatcca	taatctgggt	ctaactgtac	tttagctgaa	ggtgctggca	ggctcctgcc	71580
agggtgctgt	aagaacactt	ctattctgtg	agaatcagag	atgatttcta	gggaaaatgg	71640
gcgagaggga	gtaagcagga	gaaacaaccc	acaggcacag	ctctcatctt	tctgccctga	71700
gagccttcc	cctgccacgt	ggttttgttt	gtttgtttgt	ttgtttgttt	cagatagggg	71760
ctcactctgt	caccagggt	ggagtgtagt	ggcaagatca	tggctcactg	aagcctcgac	71820
ctcccaggct	caagcagtc	tcccaaaatt	caaagctggg	agtgtgtgtg	ccagtgtgta	71880
tgtctaggag	ccctttttcc	tgccagcccc	tacggggatt	gatgactctc	aaatgcttca	71940
gggtgtgacat	gggcacagca	gtgagtcatt	cctctgacat	tctttgggaa	gaacattttc	72000
catccaggct	tccaggcata	agatccagtc	ctctgggtgat	aaggagtcca	cagacaggac	72060
aatgtctgag	tgtatcttaa	accaggacc	atggcttgtg	ttcacaccag	accctccagg	72120
gatttttgagg	tgttttgttt	gtttgtttgt	ttgtttgttt	gttttttgag	acagagtctc	72180
tctctgtcgc	caggctggag	tgcaaggcca	cgatctcagc	tcaactgcaac	cctgcctcc	72240
cgttcaagc	gattctcctg	tctcagcctc	ctgagtagct	gggactacag	gtgtgcacca	72300
ccacacccgg	ctaatttttg	tatttttaat	agagactgtg	tttcaccatg	ttggacagga	72360
tggctctgat	ctcttgacct	cgtgatccct	ccgcctcggc	ctcccaaaat	actgggatta	72420
caggcatgag	ccaccgtggc	ccgcccatt	ttgagttttt	atgttcta	cccaaacatc	72480
tgtctacagg	cccctcagca	tattcttctc	tgggtccagt	gtcacctccc	aggcctgcag	72540
gctggctaga	gcagtagggg	gtgtgggaaa	gctctgggct	ttgcaggcac	tgatcagctg	72600
tgtgacctta	accacctga	acctcagttt	cctcacctgt	aatggaaata	ggtaccacgg	72660
cagtttggtg	caaggactag	agagtaacct	tgggaataaa	aggtagcagc	agcttgggct	72720
ctggagatgg	actgtccaag	accaacttcc	agttcctccc	cacacaagct	ctggcactta	72780
gattctgggt	acctccgctg	cttcatctgt	aaaatggagt	aacaatagga	atactttata	72840
gagttgtaag	gattgagtgg	ctggatgaac	gtcaagcact	tcaaaggggg	cctggcatgt	72900
agtgagtgat	caatataaac	cacctggctt	gtagcagggtg	tgtgtgtgtg	ggctgcaggt	72960
gttattagta	acatctgtgt	gcccttcaga	gcgtgcacca	cacttcacac	cttgtggagt	73020
ctggaatgcc	actattatag	ttcaggatag	aaaacctccc	tgcaagcact	cgcttttagct	73080
tgtctccacc	gaacaaaaca	acacaagttc	tttattactt	ggaatgggaa	aacttcaaag	73140
gcaaaaaaaa	aaaaagactt	tcgagttacc	ccaaatctta	agccaaagtc	aatgaaaaat	73200
atcaatcttc	atattcaatt	tttgcgatac	ttttgtctcc	ccagcagtca	atggagagaa	73260
tccaagcaca	cagaaatgtc	aattaccagg	ggcagggcta	tgaattcctt	tcagagccct	73320
gggctgggga	agagtgcagg	cagacagatc	tgggtcctgt	tatcacgttc	ttagattggg	73380
tgtccttgta	ggagtcatga	agcatcttag	tgcctttgtt	tgtacctat	aatgcctacc	73440
tcagagagta	ataaggataa	gtaaggctct	acgtgaaaag	tgtctggccc	tggcacatag	73500
taggtccttc	attaatggca	gctactaatt	tttattacat	acgcaaaatc	acattacagg	73560
tcaagtacgc	tacatgacag	tgaacagtt	ttttgtttg	tttgttttga	gacagagtct	73620
cgctctgtca	cccaggctgg	agtgcagtgg	cacgatcttg	gctcaccgca	acttctgcct	73680
tcaagcaatt	ctcctgtctc	agcctcccga	gtagctggga	ttacaggcat	gtgccaccac	73740
ggcagcta	tttttttggt	atttttagta	gagacgggg	ttcaccatat	tggccagact	73800
ggtctcaaac	tectgacctt	gtgatctgcc	caactcagac	tcccaaagt	ctgggattac	73860

tgccatgagc	caccgcacct	ggctgtgaaa	cagttttatt	gtgttttctgt	ggaatgtgtc	73920
ctacccaacc	tatagctaac	tcctatagtt	ccctcagttc	tcagctcaga	tatcccttcc	73980
ttctctgtact	gttacctagt	actggttttc	atagcaccag	gtacctctct	ggcatagagc	74040
ttgtcacagt	tgcaagttta	tgtaccatca	taggatttta	aaaatattca	gttggtgtctt	74100
ccattaggct	ttcatttggg	aactccacgc	aggcagcagc	tgtatatattt	gtattgccta	74160
ctgtatcctg	agaactttgt	accctactta	gcacagaatg	gaggctcagt	aaatactgga	74220
catgagagag	agagagagag	agagaggaga	gggagagaga	gagagagaga	ttcaacctac	74280
aatcccagct	ctgagcttct	agttccctga	tgggtaggac	tgtgatgtgt	ctcacacggt	74340
aatgagcact	tatgcagaag	aggctcagaa	aatttctcct	catggccaac	ggaagactta	74400
gagttctttt	ccaagctcca	ccgtttgctg	gcatgcaaaa	tttggactat	cacttaagtt	74460
ttccaagcct	tgctttttct	atccctaaca	taggacaata	ttcagcattg	ttgtttgttt	74520
gttgggggca	ccatgtttca	ggcacttagt	agattattgt	accaccacat	ttcaattggt	74580
cctcctcaag	ccctgcaaca	tctgtgaggt	ggtcatcctt	aacaactcac	agatgagcaa	74640
caggagactg	gggggatgag	ggaactgcc	aggaggtcca	gcttatgggc	agcagagcca	74700
agaatggaa	cagggtcttt	tattttttta	tttttttatt	tttatttttt	aaccagggtc	74760
ttttaacatc	cgaggaccac	attctttgtg	ctttccaaat	catcacctgc	cccatgcaac	74820
ttacagggtg	agttacatta	aacaacgtat	gtaaatggct	ttgtgctagt	tattcaccac	74880
cacaggggaa	gtgagtcacg	gacaagagt	cagccgctcc	attcggatcc	tggctctgac	74940
acttacctgg	aaaatgactt	aaccattccc	aggatcagct	gtttgtctgt	aatttaggta	75000
gtttaatggc	acttgtgtcc	tagagtgttt	tagaaggttg	aataatatgg	agcacttaac	75060
atacttagca	cctagaaaaca	cttccataaat	attagtgtct	gctgttgtaa	tcgttattaa	75120
aattttctgcc	taagatctca	tttcaggagg	cccaactcaa	tctttgacaa	gcttaaacaa	75180
aaattgcttt	tcttcatttta	ttcacttaca	cagcaaacat	gaattgagcc	tgtactgtgt	75240
ttccagaact	gtgcaggacc	agagaggcac	agggtgaagga	agcaaggctc	tggctctact	75300
ggggaaacag	caagaagatt	gctacaatga	ggtgggaaga	gggctggact	agagagaagc	75360
cctgattagt	gtccttgcta	cctttctctg	ggagagccaa	ggcaggcttc	ctggaagagg	75420
tgatccttgg	ctgaaacttc	gatgaagaaa	aggaaagagc	gcagtgggta	gggaggaaag	75480
ggcattctgg	gcagatgaaa	tgacatgtga	caaaaatattg	gtgatcannn	nnnnnnnnnn	75540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	75600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnngca	ctccagcctg	ggtcactgag	tgagagacc	75660
tgctcctcaaaa	aaatttaaaaa	aaaaaaagtc	agaagaacat	ttgggtctca	ctctgtggcc	75720
caggctggag	tatagtggca	caatcatagc	tcactgcaca	ttcaaactcc	tggcctcaag	75780
tgatcctcct	gccttagcct	tgaataaagc	ttggattaca	gatgagccac	cacaccagc	75840
cagaccatta	ttcataatag	ccaaaatgtg	aaaacaaccc	aaatctccat	caactgacaa	75900
atggataaat	agaatgggtg	ttgatccata	caatggagta	tttactcagc	aataaaaaa	75960
agtcctgata	catgctacaa	ggatgaacct	cgaaaacatt	atgctaagt	aaagcagcca	76020
atcacaaaag	gctacatatt	acaagattcc	atttaaatga	aatgttcaga	ctagtgtctg	76080
ctaactttta	tcacaggcaa	agctatgaca	ggaaatagat	gagtgggtgc	acttttgggg	76140
ggggcagagg	tgggggtgag	gcgagtgagt	actgctaattg	gtacagagtt	acttttgggg	76200
ataaagaaac	tggtctgaaa	tggaactctg	tgatgggtgc	actactctga	acatactaaa	76260
actgtttaat	tataactttg	aaatgggtga	cttgtgaggc	atggaaatta	tatcttaata	76320
aagctgtttt	acataatttta	catattttaa	aatgcagggtg	gagggatgag	ccctctaag	76380
agaagcagga	gtttgaggag	gttctaataa	ttgtgtgggtg	ggtactgagg	catataaatt	76440
tgtcagacct	catcaaaatg	tatgatgtaa	tcttcaagaa	agttgatttt	aaagaaacac	76500
caccagcacc	agggtggagaa	ggcaggaaga	agttacacaa	ggggtaggcc	aagagtgggtg	76560
gctcatgtct	ataatcccag	cactgcggga	ggccgagctg	ggtgggtgac	ttgaggtcag	76620
gtgttcgaga	ccagcctggc	caacatgggtg	aaagcccgtc	tctactaaaa	atacaaaaat	76680
tagccagacg	tgctcgcgtg	aacccagggg	gagaaggttg	cagtgaagca	agatcatgcc	76740
aatgcactcc	agcctgggtg	acagagtggg	actctgtctc	aaaaaaaaaa	aaaaaagtta	76800
cataggggac	agtggcaggt	gtcaagggca	ggcagggctc	ctcctatctc	caggataaac	76860
tcatagggga	cttagatgcc	atgtgggtcc	ctaataagccc	tccacttgggt	tcttgagccc	76920
actcttatgt	gtatcatttc	atgtcaggcc	tcttcttccc	aacccaccca	gccatcccag	76980
cctggctgcc	aacccacact	cctccagccc	ctgtcacccc	ataattgggg	ccaggaggga	77040
tgggagagtc	gccatctctc	ggtgccatct	gttgcatctt	tacagataac	catggctgga	77100
tgccgagat	cctgggggtg	agcagccgct	gttcagagca	gtgatcaaga	cctccccatc	77160
tccacccctc	aaggaatcgg	ttttcttcca	tagccacatc	agggtgctgtg	caggaaggag	77220
ttgaaacgag	aagccaggag	caacgagaag	gacactaaca	tttattaagc	actgcagact	77280
ctcacagcac	tcccacggaa	tcgatattat	tatccccatt	ctgaagacca	ggcaactgaa	77340
gctcaatgtt	taaggaaactc	accgaagtca	ccaactgata	aaagtgatgg	aagctgggat	77400
tcaaatccaa	gctaaacttc	cttccaagct	tactccacaa	cacagaggtt	ggggaaagg	77460
gataaaaaga	gaggggagcc	caattccatt	tccacccagc	tcctgaggcg	gagcttgtca	77520
gcacagctct	ctccttccca	gaataggaag	ataccatcca	gaggcaagtc	ctagacacca	77580
gcagtgttaa	ctccctgccc	caaggcagct	gcagacagcc	tatggctgta	gttactgtct	77640
ccaaagagtg	ttagaattcc	cactcccagc	ttcggggcca	ctcacacaag	gtgattgaag	77700

tggaaccag	agactctcca	caatgccctc	ctagagtaaa	tgaggctatg	taactttgtc	77760
caaatgagta	atttgaaaac	ctgggggctc	ccagctcctg	aaaagggaag	gatgtggggc	77820
cctttatatt	catactccac	tttgtgcagc	tctcccttgt	cttatgatag	ccctatttaag	77880
aaattcctct	cccagcacgt	ctccttcaaa	gagctctaga	cctgaggctg	tcagaggctt	77940
aggactctgc	ctattagtcc	cagggtctgg	atgaccagca	ggacacctgg	cattcagtga	78000
ccactggatt	agataaatga	aacagtgggc	gctgacctg	ccaatctccc	cctgaagttt	78060
gaagaggteg	agaagtgagg	ctgtccaact	tgtgccact	ctttctgtcc	acctggccac	78120
ctaaccctttt	ctggcttcca	cctgccccct	tgccatccct	ccccccagcc	caccagagcc	78180
atcttcaggc	atacctgggc	acgtgctgga	atagaagccc	tcgttcttca	gaatgatcaa	78240
caggagagccc	cagcccagga	gtacagcaga	gaagaagagg	ttctccagca	cagccgtgca	78300
ggccatccac	cagcgcctcc	ggtacgcctg	ttgcagcgtg	ggggccatgc	tggccccgag	78360
cctgcacaga	aacagagcgc	tgggtgaagg	gccccaggt	ggccccaggg	aagggtcctg	78420
catcatgggtg	gcacccgaga	cctctcgggc	cagcccgcga	ggagccccct	atggaggccc	78480
catagagccc	tgggcttccc	agccggtgcc	aaggagctgg	ctccgcgcgc	actagcagtg	78540
ccagaggtgc	acgcggcacg	gggctccccg	tgagccacta	tcggaaacaa	ggaaggtcct	78600
gtctgcgcgc	tgcagcttcc	tagcaggctg	ccgggttctc	tcacccaggg	cagggcgctc	78660
agggccgggc	tgtctggggag	aaagtccgca	tctgcccagg	tccccagagg	acagcaaggg	78720
gcagagcgcg	ctctgaagca	ccgcggggccc	atgtccggac	tctcgcgcca	ggaaagaccc	78780
ctagaagctg	gcaggaagaa	gggcaagttc	aaggctaccc	tacgacccca	tcttccagtt	78840
gccccctcaa	gacctctcct	tccctctggg	gccggggcgac	agcaagccct	ccccctttcc	78900
gtatcaggtg	acccacgacc	ctacagctct	tcgggccaag	ccaacagctg	ccacgtggag	78960
ggagaccag	gacgggctct	cctcggttcc	ctctcccccc	gcgcgcccc	cactcactcc	79020
gcaggtctcg	gggcaaccagg	ctttgcacct	cggaacccgc	ttgccccct	ccagccccgg	79080
gagggggctc	ggacttcggc	aggaagtctg	gcggtctgtg	actttataag	ggcagcggtg	79140
gcggtatggc	tggcggggcg	gtgtgtttac	caaagggagg	gaaagagccc	cagctcccc	79200
cgccgcggcc	gctgcagcct	cggcgggagg	agagggaaag	cgggcagcgc	gggggagggg	79260
agcgacaact	gggatgagac	cgaggaaagc	ggagaggaga	agggcaagaa	agacccagag	79320
agaggggagg	aagtaccagt	cacttcttcc	ggtattctca	ggtattctca	tctgtgaaac	79380
ggggctttgg	gttcaagcgc	tccaggaggt	ccgctggaac	tctggcaaac	gcgcagctct	79440
aagcagagga	agtgcagcga	gcggggaccc	gggaggaaga	gaagagtcgg	aggggtcaga	79500
gaaaagaaaa	gggaaggacg	cgcttgccga	gatgggacac	tgtgccgcgg	gaccgcgggc	79560
gcaagtaacg	gtctttcctt	gggaagcctg	gcagtgtcgg	cgggagccgg	cctcggtgct	79620
tctcagccga	cgcatagccg	gcagccagca	gcgcgcccc	tccccgccca	cgctgctcac	79680
ctccggtcac	cggcaaatga	gcagccagca	gctgcggacg	cctccgggag	cgcaacgctt	79740
tcgcggcgcg	tccggagtcc	cgtgggccc	gccctgagcc	gcgcggcgcg	tggggtcttc	79800
tctgcgtgca	ggacccggcc	gccacggagc	ttcagcctga	cagcccggtg	gcctcgccct	79860
cgctgtctcc	tcggaagaag	cgggggaact	gggaacccgc	cgggcgccag	aggtctgcga	79920
agctgggctt	ggatgaagtg	gatctgcgga	gttgatagtt	gtatttacac	gcgtccggag	79980
ctgcgccccg	aggtgggggc	gggggtcccc	ttcttttccc	ctccccctag	gtcagagttt	80040
acgcgcacgt	gactcgcccg	ctggtccccg	acactctccc	tctggcacag	cccagcacc	80100
tacattttcca	ccctggaccc	ccatcttctc	ccccaaagccc	ccagactaac	atcaggcagc	80160
gccctctgta	tccttggtca	aaacaaagtg	cgattcggtc	gaagccgact	gaccgcgatt	80220
caggggccgc	ttgggtgggg	ttttgaactg	tgacgtggga	agcagtggtt	tccgagaggc	80280
agagtggcac	gggtttcttt	ggagttagtc	agatcgaggt	ctgagtcctg	actttttaac	80340
tgactaccct	gggttaccta	gggcaagtta	cctctctgag	cctcagcttc	ctcctcttta	80400
aattcggtta	aaatggaacc	tacctaaactg	cccaaaggaa	tcgcgattgt	gatgcaggta	80460
aaatgctaag	catagcattt	ggcatagtaa	gcataatggt	aattggtgct	gctgtcatta	80520
tttcagaaga	cctggtgate	ggatgcttcc	agatcaacaa	ttgattgact	ccaggtaaat	80580
ctctcagcct	ccctgagcct	cagtatcctc	atctgtaaaa	tagactacta	tgggtgtggag	80640
taatgagaag	taatctcatt	acatgtgagt	tttaattgtgt	gttaagagtg	ctgctaattgc	80700
atgctgagct	taatacctag	gtgatgggtt	gataggtgca	ataaaccacc	atggcataca	80760
tttacctacg	taacaaacct	gcacattctg	cacatgtacc	ccagaactta	aaataaaaaat	80820
aaaatttttt	taaaaaaaga	gtgatactgg	tggccagggtg	tggtggttca	tgcctgtaat	80880
cccagaactt	tgggaggcca	aggcaggagg	atcgcttgag	ctcaggagtt	cgagaccaac	80940
ctggacaaca	tggtgaaacc	ccgtctctac	aaaaaagaaa	aaaaaatagc	caggcatggt	81000
ggtgtgcacc	tgcagtctca	gctacccagc	aggctgaagt	gggaggatca	ctgagctgga	81060
gagatggagg	ctgcagttag	ccaagatcat	gccactacac	tccagcctgg	gtgacagagt	81120
aagactctgt	ctcaaaaaaca	aaacaagaat	gactacagaa	agctccaaga	agccctcaga	81180
taaaagggaa	cccctgaaca	gatgagccac	caagccaaga	gaggaactaa	tggctaccat	81240
agacagggca	ctttccaaaa	taaaaaatact	gttattaatt	cctcaagaca	tcatggtccc	81300
atttaaacct	catagctttt	cacagaggga	gaaactgcag	gcttgaagct	ggagcaaggt	81360
tagaggtagg	atgcagagtc	aggctcgccct	ggcattttaag	tacggctcct	tccattcctc	81420
ccagaaggag	aatggcaaga	gcaaaggcct	agctgtggga	atggcacaag	gagttctcgg	81480
tggccaaagc	acatgtcagg	ctctgatggt	ttaacttctt	aaaatgcaat	actgcctccc	81540

agaacttcca	gatcaaggtc	aaactcctca	gctctacaca	gggggaccta	gagtcaactt	81600
tctaagctag	gagagtcag	gatccctttg	agaatacaaa	agacagtggg	cgcggtggca	81660
gtggctcatg	cctgtaatcc	caacattttg	ggaggctgag	gcaggaggat	cacttgagcc	81720
caggagttca	agacctgctt	ggtcaacata	gtgagacccc	tatttctaca	aaaaattcag	81780
ctgagcatgg	tggcatgtgc	ctgtagtctc	agttactggg	gaggctgaag	taggatgatc	81840
cctgagcctg	ggaggtccag	gaagctggag	tgagccgaca	tctcgccact	gcactccagc	81900
ctgggtgaca	gagaccctgt	ctcaaaaaaa	aaaaaaaaaa	gaagaaatat	gttattgatc	81960
tactcttgac	aaaaatgctt	gtgtgaacat	ggacacacac	actcatcaac	attcacattt	82020
caaggttttc	atggaccctt	tccatgagge	tctagtggtc	catggacccc	catggctgga	82080
acacttgctc	ttcctcatct	caaccacat	ttccatggag	ttggactgtc	tgctgcatga	82140
ggacacaggc	ctcatttggg	gtgttcattc	actgtgtgt	atcccagcac	ccagaacagc	82200
acctcaccta	aggggcactc	agcacatgtg	cagtgaagag	tcagtcagct	ggtttcacac	82260
ctcccagttc	ttgcacctgc	tattccttct	tgtgggaatg	acagatttcc	ttcatttctt	82320
tttttttttt	ttttgacaga	ttccagctct	gttgcccag	ttggagtaca	gtggcacgat	82380
ctcagctcac	tgcaacctct	gcctcccagg	ttcaagcaat	tctcatgcct	cagctcccca	82440
agtagctggg	attacaggtg	cacaccacca	cctgtgagct	gatatttttt	tcttttcttt	82500
tcttttttcc	tgagacagag	tctcactctg	ttgcccagge	tgaggtgcag	tggcgtgatc	82560
tcggctcact	gcaagctcca	cctcccgggt	tcaagtgatt	ctcctgcctc	agcctcccaa	82620
gtagctgaga	ctacaggcgc	gcaccacat	gcctggctaa	tttttgtatt	ttttagtaga	82680
ggcgggggtt	caccatattg	gacaggctgg	tctcgaactc	ctgacctcgt	gatccgcca	82740
cgttggcctc	ccaaggtgct	gagattacag	gtgtgagcca	ctgcaactcg	ccattttttg	82800
tatttttttt	gtagagatgg	ggtttcacca	ttgtggccag	gctggctctg	aactcttggc	82860
ctcagctgat	ccaccacact	tggccaccca	aagtgttggg	attacaggca	tgaaccactg	82920
cgctcagcct	ccttcttcat	ttctaattgt	ctcctccttc	acaactcagc	tcaagtttca	82980
cttctctctg	gaagctctac	tctaggctgg	attcaggggc	ttgtccacat	acccacaaa	83040
tactctgctt	acctctatgg	aagtccccac	actgatctag	aataatcagc	ttagttttct	83100
gcccccatcc	cgccccatga	gatgtacatc	ttgtgggggc	aggaaccacc	acgtggtagg	83160
tgatttgggt	gcctgctgcc	tatcacaggg	cctggcgcc	aataagcttg	cgcccaacat	83220
ttgttgaata	aatgaaaagg	gaatggtggg	aaagggaagc	gaaaaggtag	gctaaaatca	83280
gttttgaatt	acctctggga	ggccaaggac	tttcagttct	gcagggttag	taacaggaaa	83340
ctcctggatt	ttgttttctt	ttggttttgt	ttgtttttta	tgaagggtag	cgttatcgtc	83400
aggtttttgt	gtttaattaa	tgagcatat	attggaaaag	acagagacct	taaagcagtt	83460
aggagaccac	cataatagtt	cacattttgc	agccataaaa	aggaatgagg	ccaggcatgg	83520
tggtcactc	ctgtaatctt	atcacttcgg	gaggttgagg	caggcggtac	acctgaggtc	83580
aggagtttga	gaccagcctc	accaacatgg	agaaacccca	tctctactaa	aaatacaaaa	83640
ttatccagge	gtggtggtac	atgcctgtaa	tcccagctac	tcaggaggct	gaggcaggag	83700
aatagcttga	atctggggagg	cagaggttgc	ggtgagccga	gatcgtgcca	ttgcatggca	83760
ggtacatgga	tgaagctgga	agccatcatc	ctcagcaaac	taacacagga	acagaaaacc	83820
aaacaccgca	tgttctcact	cataagtagg	agctgaacat	tgaaaacaca	tggacacaga	83880
ggggaacatc	acacactagg	gcccgttggg	gagtgggggt	tgggggggtaa	ggggagggaa	83940
cttagaggac	gggacaatag	gtgcagcaaa	ccaccatgac	acacgtatac	atatgtgaca	84000
aacctgcaca	ttctgcacat	ggatcctgtt	ttgttttaag	aagaaataaa	gaaaaaacca	84060
agaagaaaca	aacaaacaaa	aataattccc	attttaaaca	ataaaaaata	ggccaggcat	84120
ggtgactcag	gtctataatc	ccaacacttt	gggaggccaa	cgcgggcaga	tctcttgagc	84180
ccaggagttc	aaggccagcc	tgggcaacat	ggcaaaaccc	tgtctctaca	aaaaatataa	84240
aacaaacaaa	caaaatagcc	aggagtgggt	gtgcatgcct	gtcatcccag	ctactcaggt	84300
ggctgaggtg	ggagaatcac	ttaagcctgg	gaggcggagg	tagcagtga	ctgagatcgt	84360
gccactgcac	tccacctgga	gcaacagagc	aagattttgt	ctctaaataa	ataaataaaa	84420
taataaaaaa	cagagaagag	gaaagacacc	tgagatatat	ttccatatct	gaatcaatag	84480
gatttatcaa	cgttctcctc	tacccccaaa	actaatcct	tcctaaactc	tgttctcctg	84540
acactactca	taggttaagt	ataacagcat	tatcacattg	gctgtcatgt	gggtccttgg	84600
ctagaggctg	cttcacagct	taatggacaa	gagcactgag	acagggtggg	tctaaatcct	84660
ggctctgcag	ctgattattt	gtgtgatttt	gtccaaatca	ctccatctca	tgagcctcac	84720
tcttctagtc	tgtaaatgac	tgaaaaataa	agtatccaat	tcaattcatt	atttaatgaa	84780
ttattttgac	taacaaatag	ctattataaa	tatttaggct	gggcacagtg	gctcacgcct	84840
gtaatccag	cactttggga	ggccaagggt	ggcagatcac	ctgagtcagg	agtttgagac	84900
cagcctgacc	aacatggtga	aaccccgctc	ctactaaaaa	tacaaaaatt	agctgggtgt	84960
ggtggcatgt	gcctgtaatc	ccagctactc	aggaggctga	ggcaggagaa	cgcttgaacc	85020
caggagacag	aggctgcagt	gagccaagat	cgtgccactg	cactctagcc	tgagcaacag	85080
agcaagactc	tgtctcaaaa	aaaaaaaaaa	aatctctgca	tgaagaatgt	acataaaatg	85140
gtgcagccat	ttcgaaaac	agtttggcag	gtcctcaaat	agttaaacat	agagttacca	85200
ctatagccca	gcaattccac	tcctaaatat	actacaccca	agagaattga	gaatatttgt	85260
taacacaaaa	atgtgtatc	aagtatttat	agctgtatta	ttcattacag	ctaaaaagtg	85320
caaacatccc	agcagtcctc	cagctgatga	acggagaaac	aaaatgtggg	atacccatac	85380

aatgtcatat	tatttggcca	taaaaaggaa	gtactgatac	atgctacaac	atggatgaac	85440
cttgataatg	ttattctaag	tgaagaaac	cagacacaaa	agaccacata	ttgtatgact	85500
gcatttatat	gaagtgccca	gaataggcaa	atccacagag	acagaaagta	gatttagtgg	85560
tgccagagac	tggagggagg	agataatggg	aaatgtggaa	tgactgctaa	tggtatggg	85620
tttcttcttg	gggtaatgaa	aatgtttgtac	aattagataa	tggtgatcat	tgtaaaactt	85680
tgtgaatata	caacatgctg	aattttatac	tttattatat	tttatttttt	ttgagacaag	85740
gtctcgctct	gtcaccaggg	ctggagtgc	gtggcacgat	ctcagctcac	tgcaatctct	85800
ctgcctccca	ggctcaagca	atcctcctgc	ctcagcctcc	tgagtacctg	acactacagc	85860
atgtgctacc	atgcctggac	aatttttgca	tttttagtag	agacaggggt	tcgctatggt	85920
gcccaggctg	attttgaact	cctggactca	agtgctccgc	ccacctcagc	ctcccaaagt	85980
gctaggatta	caggtgtaag	ccaccactcc	cggcctaaat	tgtattcttt	aaaagactga	86040
attgtatgg	gtgcgaatta	tatctcaatt	taaaaaaaac	aaaacaaaac	aaaaaaaaaa	86100
cctttgctg	tgtcaggcac	tagggattcg	atgctgaata	agacacagac	cctaccctca	86160
gagaacacag	agcccagcag	gagagagtca	cagatgaatc	aagtgttaca	tcattctatag	86220
gaagcgccat	ggaagaaaga	catggtgcc	tgagaacata	cgcttagaga	agggaatttc	86280
atctagactg	gggctcagg	aggaatcttt	cagggtgatg	cttgtgctca	gagttttcca	86340
tgtcagaatc	agtagaattt	atcaatcctc	cagaggagga	aacagcaa	gaaaaatctt	86400
acaacaggag	gatgcggaga	cattccgaga	gctgatcaag	ggctgggtgtg	aacaaagcac	86460
ataggatgca	gagcctgtg	tgtgaggttg	cagctggaaa	ggtaaaacac	taattacatt	86520
ggatcttctg	agacaataaa	gagtatgcaa	taattctcaa	cgaccgaaac	tgaccttctt	86580
cctccctaac	ttgcttgctt	ccactgttgc	ccgtatcata	aaagcaccac	cctcttctac	86640
ccagtggctt	aagacacgaa	actcaagtca	tcccaggctt	tctccccacc	tcactctcca	86700
catccagcct	atcagcgagc	ttgtgggtct	taccacgtaa	agacttctca	tctccagcta	86760
ctaccatccc	ccaagcccag	atcaccatca	gctcaggcct	ggactcctgc	aacctttcta	86820
accgggtctt	cccaatccta	cccccgcaac	atgaccccaa	tagcccatca	gaatggacta	86880
atcgagatgt	agattttgatc	aggccacatc	ccttgaaagg	cttcctgtga	ccctcgggga	86940
aatgcaaaaa	ctcccaatga	tggccctgga	gtcctgtgce	atctgggtct	gccctctgcc	87000
ctctgtgtct	ttgccatggt	aacctccttc	acacccatta	atactccatg	ctctctccta	87060
cctcaagtct	ttcctgggct	ggaacattct	ctgcactagc	ctagccaact	aaccttttag	87120
atcttttgtt	tgtttgtttg	tttgtttgtt	tgtttttgag	acagtcttgc	tctgttgcca	87180
ggctggagtg	caatggtgca	atctatctcg	gctcactgca	acctctgcct	gccgggttca	87240
agcaattctt	ctgccttagc	gtcctgagta	gtcagagcta	taggcacctca	ccatcacgcc	87300
cggtaatttt	ttgtattttt	agtggaggtg	ggttttccacc	atgttggcca	ggctgggtctc	87360
gaactcctgg	cctcaaatga	ccaccctcct	cggcctccta	aagtgtctgg	attacaagca	87420
tgagccactg	tgcccaggca	acacttcaga	tcttaatgat	catttctctt	aagtgcctga	87480
cctctttag	taactagcct	gactccagca	atgaatcctt	ttgcaatgta	acctatataa	87540
catctgagtt	tccttttgat	aaaactcatc	atataattgt	tcctctgaca	gttcagagg	87600
caaggcctt	tgcccacctt	cctcaccact	atcctctcac	cacttaacac	agaactcacc	87660
accaccatg	cctcctgcct	gacaaattcc	taaccatcct	tcaaatctca	ctcacctatt	87720
accttctggg	aggcagtctt	ccctgagcac	caagacaatg	ggacacattc	ctttatacac	87780
cctgctgaac	atctcttttt	tgagggcg	gtagagatga	gtgtctcact	atgctgcca	87840
ggctgacctc	aaactcctgg	cctcaagcga	tcctctgcc	ttggcctccc	aaaatgctgg	87900
gattacaggc	atgagccact	gtacctgacc	gcaactgggt	tagnnnnnnn	nnnnnnnnnn	87960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88020
nnnnnnnnnn	nnnnnnnnnn	nnnctgggc	aacaaagacc	cttcctctac	aaaaaaaaaa	88080
aaaaaaaaaa	aaaaaaaaaa	aaattatttt	aaattagcca	gacatggtag	tgcatgcctg	88140
tagtccaagc	tacttgggag	gctgaaggga	gaggatcact	tgagcccagg	aggttgaggc	88200
tgcatgagc	cgtgatcgta	ccactgtact	ccagcttggg	caacagagtg	agacctcatc	88260
cctaaaaata	aagaagaaaa	tatggcaatt	tgactgtaca	tctctaattg	gatatacct	88320
aaggatgaga	aaggaataag	gaaggacaga	aaaaaggaaa	caaagaagta	gcaacagtat	88380
ttagcaattg	tattgttatc	aagtaacatc	aatattggta	aaaccagtaa	ttatatattaa	88440
aatactatat	atgtgtatgt	acattttacat	atgcatatgt	taggaaccaa	gtttatcaga	88500
ggaagagaaa	gggctacaaa	tgtaaaatca	aggaaataaa	aatttgaata	aaaatatcag	88560
tattacagat	ttatgatatt	tttcttataa	aaaaattata	tatatgttaa	ctctatccaa	88620
aacccaaaag	cagtgacaac	ccaggagcaa	taaaaaacct	cagcatccag	actgtagtct	88680
ctaccatttc	caattaaaga	aaccaggggc	tagttgggaa	aatgacaat	ttcatgtcta	88740
gggcaagaaa	cacacctagt	gaaatggacc	tgaacattta	attgtgttag	aaagtaagga	88800
aactctctag	aaataatgtg	atctcatcta	aaagacacag	attctgggct	ggtaaagttt	88860
tcaatggcca	aaggtgagac	aatttgagca	tcaagaagaa	tcatgacaga	acagattaaa	88920
acatgtcaaa	tatattttaa	aatgaaatat	tataaaagaa	acaattagta	gccatccctg	88980
aaggtcacta	gggcaccaac	tcatatttca	aactggtaaa	ttaatgtgta	agccaagcat	89040
ttatttctgg	gtaacaaaat	agtaaggaat	gtttttcttt	ctagaagaat	tctagtgtat	89100
aaaagtagaa	gatagaaata	gaaaatcatc	cttttggcca	ggggcagtg	gtgtaatccc	89160
agaactttgg	gaggccaagg	caggtggatc	actggagatc	aggagtttga	gaccagcctg	89220

gccacatg	tgaaaccca	tctctactaa	aaatacaaaa	attagccagg	tagtgggtgc	89280
ctgtaatgcc	agctactcgg	gatgctgagg	caggagaatc	gcttgaacct	gggagggcga	89340
ggctcgagta	ggctgagatt	acgccactgc	actccagcct	aggcaacaca	gtgagactct	89400
gtctcaaaaa	aaaaaaaaaga	aagaaaagaa	aatcatcctt	ttgcgatcct	aatgaaataa	89460
tgggcctagg	cattgatcat	taatggctcc	taaaatcact	aagtatatgg	ttgatgggaa	89520
actttatagt	ggatggatca	gactcgcaat	gtctaaacca	gttgatcaat	cttaacatcg	89580
taacaagaca	acagacacca	ggggctgctg	acaggagaac	agaggaaacc	catagctcta	89640
ccactgagtt	attcacggca	aaaaaaaaaa	aaaaaaaaatt	aaactgcggt	tcctccaagc	89700
ttctaatacct	gttgtttaca	ggaaataccc	aaggaaagga	atacttttaa	atgacacatt	89760
aaaacaacgc	caaataccaaa	atatggggaa	atgacccagt	ttcttcaaca	aataaacaag	89820
aaaaggtagg	ggggaggact	gttctagatt	ttaaaagcta	tagaagacac	agcaacccaaa	89880
tacactgcat	ggaccaggca	tggatcctaa	ttggaacaaa	ccaactgtaa	aaggatgtat	89940
ttgaaatgat	tggagggaatt	tgaacagtga	ctgcacagta	gatgatatga	agaaattatt	90000
gttattttttt	aggtgtgatc	atgattttat	gggtgatgtt	aagtaaaaga	ggccttattt	90060
gttagagata	catgtacggg	tatagagaaa	tatttacgga	tgaaatgata	cgatgtctga	90120
gatttgcttt	gaaaactcta	gcagggtgtg	gagaagcagg	tgcatgggtg	ggggaaggga	90180
tagatgaaat	agtatgcaa	aatgttagtc	tacttttgtc	cctcctgacc	cagcagggtta	90240
aaatacctca	gcatacctct	actcctccaa	ccagggtccaa	ggatcaggcc	aaaactccct	90300
gatgtggtaa	acagcctgac	cccttcttac	ctctctctct	ccagccactt	ccctaagatt	90360
ccccagtgct	ctgtgcccta	gccagcccca	ctcatctgcc	cagattcctc	aatgtttcac	90420
tctctcattc	accattttga	ccccactgtg	ctcctgggcc	actctccaag	gccccgcctc	90480
ttcatctcct	ccctccttat	tcataccttca	ggtctgggct	taggtgccat	tcctccagg	90540
aagccttccc	tgacaccaat	cccatcctca	cctagaacag	attatgtgcg	cttctttgtg	90600
cccccatgg	ccccctgtgg	gtttgcttca	cggattataa	ctgcctgact	acctgccttt	90660
ctccaccctc	tagactgaga	acaccctgag	aaaaagaaca	catctatctt	gtctgtcatt	90720
gaatccctgg	tgtctggcac	catggctgac	acataactat	cactcagtga	ctagtgtttt	90780
aatgaatgaa	tgagtgaac	tagacaggg	taagaacaaa	agagaagacc	aggcatgggtg	90840
cttcacgcct	gttatccag	cattttggga	ggctgaggcg	ggcagatcac	ccgaggtcag	90900
gagttcaaga	ccagcctgac	atggtgaaac	cccgtctcta	ctaaaaagac	aaaaattagt	90960
gggggatgg	ggcacacgcc	tgtaatccca	gctacttggg	aggctgaggc	aggagaatct	91020
cttgaaccca	ggaggtgcag	gctgcaatga	tctgagatca	caccactgca	ctccagcctg	91080
ggcaacagag	taagactcta	tcacaaaaaa	aaaaaaaaaa	gagcgaagaga	gagcgaagaga	91140
agatgtcatg	gggtaaatga	agacctccct	tccgtgttcc	ctgaccagcc	cctgccctcc	91200
cccgcactct	acctgtcttt	cttgcttcc	tctggtactt	ctgtttaagc	cgttccatga	91260
gcaggccatt	ccagggggca	cacagcactc	cgaactgagt	gaaggcaaa	gcatttgtgt	91320
aggtgctgac	tgcagaagga	agagagaggt	tggltgatga	gaagtttcca	aaactccctt	91380
ccaggcagg	actctccac	cttacccttg	tctgcatgtc	cctcctcccc	acaccatcag	91440
acctcctct	ggtgtgtaca	gccctgctgg	gaggctctgt	gttcccagct	gggacatgca	91500
gatgggctac	ctcccagccc	taccacatac	ctcgtgccat	gtccccaccg	gccatgttgg	91560
tcagcaagga	gttgagagtg	ccaatgaaga	ggtagtgcga	caactgtatc	acagacagcc	91620
acaccaggtg	ccaggcaaa	cgccgagaga	aagcgtagct	ccagaaggag	cggagttcct	91680
gcttctgccc	tgcccctggg	gtctctaatt	gggagaggag	gatctgggag	tgaattacga	91740
ggaaagtgga	caggtaggat	ggggagtgtg	gaggcttcaa	tggaacattt	cagatccggg	91800
cccaccttct	accttggct	cccagaactc	acccggcggt	ctctgcacct	gcctcggact	91860
cacacatccc	cacctccctg	ctttgtcatg	ctggccctac	caccttggat	gacctctgt	91920
gttcttctct	attgaaatcc	gatctgtctc	tcacagcctg	gtcaatgaca	cttcttgcag	91980
taataccttc	ctgatctttc	tcagcgagaa	aggtgaaagg	aacgacaagg	agaggagaaa	92040
gtcagaagg	agaggagaat	gagtgtggat	actctgttct	aacctgctcc	tcagcacctc	92100
cctttctttt	gataccagta	tcctgagttt	ctttgggaaa	tcttctctta	ccctaattctt	92160
catggtccag	atgggacat	gaattcagtg	ttctgttctt	ctctcaaggt	taaccaatga	92220
gatggttctt	cctaacaagg	cagaccggcc	atgagtttag	attaggatgg	acttaattcta	92280
aaataggtcc	tacacctgg	caagttcaat	gtcctccctt	gattttggaa	gcttcccaga	92340
acctattctt	tctttcttta	aaataaataa	ataaatacat	gttttggatc	caattgtcag	92400
atggtaaaaa	taaaaacaaa	aaaatcaatt	ttattctgta	tatttaagat	atacaatatg	92460
aggtcatagg	atacatatag	ctactaagat	ggttactaca	gttaagcaaa	ttaacatatc	92520
catcatctca	catttctacc	tgttttgtga	caagagcagt	taaaatctac	ttgttttagga	92580
aagtcccaaa	cacaatgcag	tttgatgacc	tacagtcttc	gtgctgtgca	ttagatctct	92640
aggcctgttc	atcctgctca	tctgctcctt	tttgccttc	gacctgcac	tcccatctcc	92700
tctcccaccc	cgttcttatt	tctactgtag	ctagctgcgg	tttgtgatgt	gtgtaaccaa	92760
agacgcagaa	cagagaggaa	ggaaaggaa	cagtgataga	gttgggacaa	taagagagg	92820
cggacccagg	agacctggag	aaatgggggc	actgtaccag	acttagtgca	atggcatcac	92880
agaagagggc	agaaccgagg	agtgggggga	agggaaaggca	acccatggca	ggcgggcttc	92940
aaggggtggg	gaagtgatag	gatgcgaaat	agagaaaaga	gggacagaaa	agagacgaaa	93000
gccctggacc	ctccattaag	tgagaggggt	gggaagatgc	ctaaggccct	ttttctgtcc	93060

tgcccttccct	gattctgggt	ccctggggga	gctctggagg	tgaggggcca	ggaaaggcac	93120
aaggagaggc	ttgggtctgg	aggagagatg	ggtagccag	cagggtcac	cttccttcgc	93180
tgaaaggaac	tcctttgact	gtagctccct	gttttcatgc	tcagctgttt	ccttctcttc	93240
ctttgtggtg	ccattcccag	ggcacaggct	atggaaacaa	aagccccacc	agcaaggcca	93300
aggactgtga	gccgaacctg	agactcagac	tggagggaa	agcatggtga	atcccacatt	93360
ccaccgcact	ttggaatcac	cttttagcca	ctctgatgcc	cagggtgcag	accagaccag	93420
ttaaatcaga	atgtctggag	gtgagagcca	ggcttccttt	tctaagatct	ctatgtgaat	93480
ctagtgatcc	taataagcag	caaagtttag	gaagcatgaa	aagagtaggg	caggcccagg	93540
ttcaaatecc	agctctgcct	cttcctagca	acagaaagat	ggctcagaat	taacccttct	93600
gagcctcatt	ttttgcattt	agaaaatgga	gataaggata	tctcagagga	ttattgtgag	93660
gatgaaatca	gagagcacat	ggggtctgac	aattagtaag	tgagcagcaa	aggaatgcc	93720
ttcctctact	ccttgtggca	aatgactgca	aaaatgatca	catttcttca	cctcctctgt	93780
atttccccca	atttgaatga	gactgcagct	ctatttcccc	atgccctgaa	tctggggccag	93840
ccttgtgaac	tgcttcagcc	aaaagaatgc	agcagaagtg	gctgtgcca	ttccaagctt	93900
aaatctcaag	aacgcttgtg	catttctgca	ctctttcaga	accctgaaat	cacggtgtga	93960
atgagccac	gctggcttgc	tggaggatga	cagccacgtg	accagggcat	cctgtcact	94020
ccaaacctat	gtgagtggag	ccatcctagc	atagccagcc	cccatgtaat	cctccaaatg	94080
atcagatgta	tgaatgagcc	ctgtcaaaat	catctacatc	tggccctgat	cagcggaact	94140
agccagctac	ccacagactt	gtgaaaaata	ataaatgctt	aacatttttag	gctgctgagt	94200
tttgagatag	tttgttatgc	agcaatagct	aacagatgca	ctgctccagt	cctcctcctc	94260
tctgtgata	ggtttgcctt	accctgtcca	tcccacccta	gggccaatga	ggggctctgg	94320
cccacaatca	ccagatagtc	cttaccata	gctgtagtgt	gggggcagtg	ggtagggat	94380
gtgccccgg	ggcatcagga	ggaaagtgcg	tgtacatgc	cagggtactgc	agacagagat	94440
gaagatgaag	gaggccctga	ggctgatgcc	tttttcataa	agaagctgca	gaaggagaag	94500
gaaaaagtca	gtgtcacacc	cacgttcata	gcagcactat	tcacaatagc	caaaggatgg	94560
aagcaaacta	aggggtccatc	agcagatgaa	cagctaaaca	taatgtgatc	tatacacaca	94620
atggaatatt	attcagcctt	aaaaaaggaa	agaggcaacc	atgctggctc	acacctacaa	94680
tccagacact	ttgggacgca	cgaggatcac	ctgagccag	ttcaagacca	gccttgacaa	94740
catagtga	cctcacccc	ttctctagaa	aatttttatt	taattagctg	ggtgtgggtg	94800
catacacctg	tagtgccagc	tactcaggag	gctgagtggg	aggatttctt	gagcccagaa	94860
gtttgaggct	gcagtgagtc	atgactgggc	cactgcaccc	cagcctggac	aatgaaacat	94920
gaccttgctt	ccaaataaaaa	aaaaaaaagg	aaaggaaaga	aattctgaca	catgctgcaa	94980
catggatgac	ctttaagagc	actatgcagg	gcgaagctca	gtgggtcctg	cctgtaattc	95040
tagtgcttta	gaagaccaag	acaggaggat	tgtttgagtc	caggagcttg	agaccagcct	95100
gggaaacagc	aagacctcat	ctctactaaa	aataataaaa	taaatacagct	gggctgtgat	95160
gtgcacgcct	gtaattccag	ctacttggga	ggctgagggtg	agaggatatg	attacatgat	95220
tacatgcctg	taatcccagt	actttgggag	gctgaggcaa	gcagatcacc	tgaggccagg	95280
agttccagac	cagcctggcc	aacctggtct	aaccccgctc	ctactgaaaa	tacaaaaatt	95340
agtggggcat	gggtggcacgc	acctgtaatc	ccagctactc	gggtggttaa	ggcaggagaa	95400
tcgcttgaac	ccgggaggcg	gaggttgagc	tgagccaaga	tcctgccacc	gcactccagc	95460
ctgggcaaca	gagcgagact	ctgtctcaaa	aaaaaaaaaa	ggttaagata	gtaaatttta	95520
tggttatgtat	attttattgc	atacaaaaaa	atcagcagaa	gaggcagggg	ctggaacctt	95580
gttttctaac	gagtcctagt	acaagccatc	acctactatc	ctgtaagctg	attagggaca	95640
cctggtacac	acatgcccc	acccacccca	agacacaccc	ggcagtagag	gagtcctcat	95700
acgacccatc	cccacagccg	gtggagcctc	ctcgtgtggc	tcccagaga	tcttctagcc	95760
cagtgccttt	tttcccccaa	cgacagcaaa	ggccttttgt	tcaaaagaaa	ttttacacaa	95820
aaattcatct	tacaaaacac	accaatgggg	agcttgccag	tcacttccct	ctttattctc	95880
cttggtgact	ggtatgacat	caaagagaat	ccctaagttc	ctcaacagct	cagtttgaaa	95940
accacogacc	tagcccaacc	tcctccatt	ttacagagag	tgacgttgag	gtccagagag	96000
gtgcagtga	ttgctcaata	aattgacaga	gtaagcagca	gcaaagtcag	attaaactaa	96060
gaattcctgt	tcctgtctcc	tttccccttc	caactctaga	gagacaggag	agaggctggg	96120
catggtggct	catgcctgta	atttcagcac	tttgggaggt	caaggaaggc	ggattacttg	96180
aggtcaagag	ttcaagacca	gcctggccaa	catggcgaaa	ccccatctct	actaaaaata	96240
caaaaataag	ctgactgtgg	tggcacgcct	atagtcccag	ctactcagga	ggctgaggca	96300
ggagaatttg	ttaaacccac	taggcagaga	ttgcagttag	ccaagatccc	accaatgcac	96360
tccagcctgg	gagacaaagt	gagactccaa	ctcaataata	aaaaaaaaaa	aaaaagagag	96420
aggaaagaaa	gatgaggcag	ccatctgggt	tctccagggg	aaggaggagg	aaccagaaaa	96480
gtgactctta	tgccaggagt	agaaaggctt	gagtgccctc	ggggctcagt	ctctgcataa	96540
ccctccaaac	ctccaaagct	tatgggacta	agctagactc	atgtctgggt	ggtgactgcc	96600
agagatcctc	ttctctgccc	ccataacctg	caggcagtg	caactgcctg	tgacctaaaca	96660
ctaagcccag	agagaagtcc	cagggtggat	ggcttgagat	ccacactctt	cccttctctt	96720
cactcagcca	tctgtggtgt	gctggcttta	gtcctccagc	ttgctgcctc	ataattgaag	96780
catggttgcc	acaactccag	ctatcacatc	ctcacaccac	aacattcaat	gaggagagct	96840
ttgtttttac	tctgctttca	ccttgcgctca	gggaagaaaa	gtcccttga	atcttccact	96900

atatacactc	ccttttatctc	attaaaaaagg	actggatcat	atgctgacct	ccacctatca	96960
ctagcaacgg	gtaaatggat	tgccatgggt	ggctttaatc	aatcaggatt	catcccctgg	97020
gctaagcggg	tactgccc	gataaaactg	ttcgcaatga	ataagacaga	atgggtgttg	97080
attgacctct	aatagccttg	gcaacagttc	atccccctgat	acccaacat	cagccactgg	97140
gacagctgga	caagcctctg	tgtctgcccc	tgctgtacct	actagccact	tgccaccttc	97200
ttgtccaaac	tagaagctca	cagcagcaaa	cgccccactc	taaagggtccc	ccagcctcta	97260
cccaacactg	gccaagcac	attatgacca	ctgccacaaa	agcttgggca	agtctgaaga	97320
aggggcttag	cggttacaag	ctcaggctct	agaaccgaca	agcctgggtt	caagttccag	97380
tatcatggct	actagctgca	gaaccttcaa	caagcttttt	aacctcagag	actcaaatgc	97440
ttcatctgta	aaatgggggt	aacacagtac	ctacctcacc	gagttgatgg	agacaaataa	97500
tgcaggttca	caagacaagt	gtctggcata	tacaagtgcc	cagtgaatgt	aggctgttgc	97560
tatatctacc	ttaataataa	ggaagactgc	cgaggaagag	tcaaagtctc	cattgtacag	97620
agtgatgatg	gtcgaaacgg	gttggccaaa	taggttccca	atctggggat	gataggacta	97680
gcctggatca	cttattttatt	catgaaacag	atacttcctg	agcaccacgc	atgtggcaga	97740
ccctccttat	acccaaactc	accctccacc	gctagagctc	ccacctcagc	ttggggccaa	97800
cccatctgag	gcagccaatt	atagaaaagg	gtctctcctt	ccctccacct	ccccgccacc	97860
ctgccgagtg	cctgggatta	gggaaggctc	ccacctgcag	gttgggtgatg	agaaacagga	97920
ttcccccaat	ggtgagcatt	ggcatggcca	ggaagagcag	cacggctgag	cctggaccat	97980
caaagtcaga	ggtagggtgg	tgtcatagtg	caacccaaac	atggagcccc	aaactctgct	98040
cccacctgct	ccaaattccc	aacaatcctg	gtatccaggc	cccaattcta	gccagcgttc	98100
cagcgtcctt	caaggggtttt	taggataccg	gccaaggctt	ccccagatat	ctctgtggaa	98160
gtcttctgag	ccaccttctt	cccaaccaaa	gttggtcctc	agtcgtgtgc	aggcgaggaa	98220
gtctacagac	agaggcagag	ctctaagtga	ccagacctct	ctcttccctc	agtaaaccac	98280
aagctgcctc	tccctttcat	ccttgacact	cctggaaaag	aagaccctgg	actcaggtcc	98340
ctggtcaac	cctctagccc	attcccta	tcatgggtatt	ggccttgagc	ttcaatcatc	98400
tgtttaatgg	gaacaacagt	tectgtcttt	cctgtctcag	gtgctatgag	aactgagtga	98460
gaaaaggacc	atgggtctttt	ctttgttcac	taaactctga	gcacttcttt	ggtgccaggc	98520
attgtgcttg	gcactggaaa	tgcaagatga	atcagatagt	ccttgccctt	aaatagactg	98580
acatgcaaac	aaatggttat	aacaggctctg	gtaagtgtga	gaccacagca	aaaaagctca	98640
agagctgggc	taggggaacc	cttgacaaat	tcttctccc	caaaccagac	ttctgcccac	98700
cattattctg	gccacaacct	atgcctgtcc	tattatttgc	taaaatgttt	taagttgact	98760
cacttttatc	caaaaagtat	ctatttttaa	aggacacttt	atatcactac	tgtagatgaa	98820
aacactggca	ttacttgtca	tgaatagaaa	gtaactgtca	aaataaatac	aatgaaagga	98880
aaacaatggt	attcaattgt	agctggatgc	atlttgacctt	agaatgttca	aagcctaaga	98940
cctgctcttc	ccatcagtgt	taaaatcaca	ctggccccac	atgaagacat	tctttcatga	99000
aatcagaagg	actgaaagag	aaataaaaaag	ggaatagctg	ttctaccagg	tgattttgatg	99060
tttggttagtg	tagttcacgt	agtatgcgtg	tgccccctaac	atcctcttaa	ctaccgtgct	99120
ataccttaag	aagcactgcc	aagagctaat	tttagagtat	tcacacagtt	taccattcaa	99180
tttctgtctt	tataaaatgt	acatctctcc	tactactaaa	ggttgagagac	tcctttcaca	99240
atagagtcct	tatgggtcca	atgctttttt	caaaactgaa	aagccctata	ttatggagga	99300
agaggaggat	tgttgctcag	acgatttgca	ggcacgagtc	aaacattacc	cagccaccac	99360
ctccacattc	agttgcttaa	aaatcattta	caggctttta	gagtagatga	tgctggtttg	99420
ataaggagag	tggtttgaaa	taattgggtt	gaggtgctgg	gccatctcat	gagatctgtg	99480
tgaacaaaga	cactcagcct	ctgtgtttgc	ccagcatgag	tgcagacaat	ctcatgatgc	99540
tgtcagcttt	agcatagctt	acacacacaa	gagtaatgta	ctttctttcc	taaaacaaaa	99600
attgagccac	gggtctaaca	ctaggaagga	atattgggag	gcacatcctg	gccaccataa	99660
ccaaggcaat	gacagaaaga	agagtggagg	atcaggaggc	ctgcacatca	ggcccacctc	99720
ccacttgctt	tctctgtggc	catggacatg	tctttgcaag	gggtcctgct	gtggcttcag	99780
tttctcctct	gtgtaatggg	tggaagggtg	gtggaaaata	aaccagattg	gagttccaga	99840
cttaacagac	tggtgaaatt	ttaaaacaaa	gatttttgagt	acaatagggt	tgtcaacttt	99900
taccctgcta	agtaaggata	tttgcaaaat	ggtcattcat	ataatcattt	cattaaaaag	99960
agaaagagaa	catttttaaca	cataggagaa	ggatgtaaag	gttttttgtt	gttgttttgt	100020
tttttaggggt	tttttggttg	ttttttggta	gagtctcact	ttgtcacc	ggctggagtg	100080
caatgggtgtg	atcttggtct	actgcaacct	ctgcccctcg	gggtcaggca	attctcctgc	100140
ctcagcctcc	tgagtggctg	ggattacagg	cgcgcaccac	catccagcta	atttttgtat	100200
tttttttagt	agagaaggga	tttcaccatg	ttggccaggc	tggtcttgaa	ctcctgacct	100260
cagatgatcc	acccgcctca	gcctcccaaa	gtgctgggat	tacaggagtg	agccactgca	100320
cctggccaga	tgtaaagttt	tgaataaatt	ctactctctg	aagtaatccc	tctccatcat	100380
ccttgctttt	cacattttct	caataaaactg	ttttcacaga	ccagcaatag	ctcaagatcc	100440
ttccaggatt	ctttcaagct	gcagatctct	gaataaccat	gtggctctgta	tatcttgctt	100500
atagccctct	gctcacacct	gccccagcca	ccagggtgct	ctgagcttgc	atccctccca	100560
cccacctgac	agcactcacc	tgcagagggtg	aaggctatga	tgagtgtggc	gggtggtgtag	100620
aaaaatctga	aacacataac	aggaaaagca	gaatattgtc	aaggaggggag	aaacctggga	100680
gaaaaaacat	gattctgctc	agccagccca	caagtgtagg	acttgaccgc	accctcagcc	100740

tgggatgcaa	cgggcactga	tgcctctgag	ccccaggetc	aaaaccaggc	gcaagaagcc	100800
gcgatgagat	tgagatgtgg	tcctgacctc	atggacagtg	cattttgtctc	attctgaggc	100860
ccaaggctag	catggaaagt	cttggacaat	gagctcagct	gacgatgtga	ttggcttggg	100920
acttagccag	gacagaatgg	gcaaagcgaa	ggctctccca	cctggaagcc	ccaacagccc	100980
aaccccttgg	agaaaagggg	tagtgccctg	tctgcaaatc	aaggccttga	gttctaactc	101040
ctectcactc	tgtgaccttg	ggcaaggcgc	tgtccttctc	tgggcctcag	gagccttttc	101100
tataaaaaga	aatgatcgga	ctgatctagc	tcagagtgtc	atgatttcag	gactacagtc	101160
ccaaggttat	caggctccct	tagcatttgg	gggtcttgtg	aggcatggag	taaaaaaaaa	101220
aaagcaatat	cctaaggctg	gagaagaggg	aggggacaaa	ggaaggggag	gaaaggggag	101280
gtagcaggga	gccaaggacc	aagaaggact	gaggtacagt	cattctgcat	ccaaaggctt	101340
aaatgtgaag	ggactggctt	tactctggct	gtttccggaa	aggcaggccc	agccagccct	101400
cccgtctctc	tctctgacag	ccaatctcac	atgtgcctcc	ctgggagcac	ctgctctgag	101460
ctgtatcagc	ccccagcagg	ccgctgatta	ccactgagcc	tggccacaga	gcacgagatt	101520
aggatgcagc	aacacactgt	gtgtgagatc	acgtcccga	ccttctgact	catctgcaca	101580
ggaaccccc	ccagtctccc	ctccagtcag	aagggacgtg	aaattccacc	agtggcaata	101640
ccaaagaaac	ttcctattag	ctaagccctc	aggagtgtat	tggctgttgg	ggcggggagg	101700
ggggcggtg	aggaggatga	ggatgaagcc	tgggcaacct	ggatgtgagg	ctgtgcaggg	101760
gatgaggaca	aggatccttg	gggtgaagga	agagaagagc	aatttttaggt	tttgctaatt	101820
ttgtaaccct	ggctccaagc	cagcccttac	aggaagtcac	cctggcctcc	ggctcaattc	101880
agcacgtgat	agggaaagcca	catttatgca	gagcagggaa	cgaggtaagg	aaatgggaagt	101940
ggggctgtgg	tgaagtgggc	aagtctagag	agagtcccg	tgcctggggc	gttcctaac	102000
agctgctggg	acgagctgc	aggtgtgtg	cctggcaggg	tggccgggct	gtctgactct	102060
ggatttcact	ccaagctagg	ctgctgcttg	aaggattcct	cttaccacc	tttgctggg	102120
ctggcctttg	ggacttacat	ggctatgagg	cgtgccacgg	tggctctgaa	ccggtcaaa	102180
atgtagccag	tggggaatgt	catgaagtgt	ttcatgaag	accccagggt	gaagatgagt	102240
gagaacctct	catcctgggc	tttgagctct	ggagtgaaga	aaaggtctcc	catgcatccc	102300
agccttctctg	ccaaatgagc	acacaggctg	ggctcccctc	cacctcagac	agcttgtcgg	102360
tcgcaaatct	gtcccttaag	ctgagttgaa	atgtggctgc	ccctaattac	ccctcaggag	102420
ctggtgcctc	cctcccaggc	acttcccaga	tcaagtgggg	tgagagctgc	tgaccttcc	102480
tctcatcata	gaaagagggg	tgggcagggg	gcagagtcct	tctgtctcct	tgccaccacg	102540
tgggagccag	acttaacttc	cttagaaaa	tcacccctgc	ccttaccagc	ctgcccgtg	102600
gcattgccaa	tcggcccagc	atctggtcca	cacagatcct	ttaaagtaac	ttcattcttg	102660
aagacaaaca	ctagtgaagg	ccagccaaag	aggacgccag	caaagcccag	gcattccagc	102720
agcccagtc	gcagtgtggc	cacgtgcagg	ggcaggccct	ggcccgccat	gagcagaagt	102780
ggagtggatc	ttcaaataccc	actttgtcct	cctggacgga	tcacaggcgc	cgtaagcctg	102840
gcgtttgagc	acttgaaaa	ttcctctggc	aagccaagcc	cttcctttcc	cgtagctctc	102900
tggttgtttc	aggcctgggc	aaaaaccatc	agcgggtgat	tctctggatc	ctgtagaata	102960
aagatagagg	ctgctggaag	aggaggcctg	cgggaaaggg	aaaggtagac	tagagttatt	103020
tgtgaggtgc	attaagaggc	aggatgatca	tggccgctgg	cagcaaatgt	ggggaataaa	103080
tactccaata	catcatctta	ggcactgcat	ttgagtaacc	acgtggcaag	tagagaggca	103140
ggtcttgatg	gccacctgga	gtcacagggtg	agataaacga	cttacccaat	agctccccgg	103200
gagcaggtgg	agaagcggag	ctcctgctgc	cgaattctga	ataccgtccc	ctaaaataat	103260
gacagcaact	caaccagggtg	cagaggcagg	gagatttcat	acagaagaca	caaactcccg	103320
ctgccaaagt	ggtgttatct	tcagtttact	gacaatgaaa	caaaagctcc	catggatttc	103380
aggaacttgc	ccaaggtcac	agggctagtt	tagtcacgac	gcaggccatt	ctactgccag	103440
aaataacccc	aactcccatg	accctcgcct	aggactcgca	aacctggtcc	ccgcgcctcc	103500
tcctcgcac	aacttctacc	aggaaagcct	ccgggggccc	ctccccgcca	gcctccgcac	103560
cccgtccag	cctgcggcct	gccctccccg	cagaggagcc	cgaggggcca	ggccgcgctc	103620
ggcgcccat	ggcgcccgaa	aggggaccct	tcgccctacc	cgctctgctc	gcgcccgggc	103680
tctccgcgcc	ctttccgcac	gggccagggt	cgcattcgcg	cctctcgag	ccccctccag	103740
tcccctgctc	gcctccgccc	cctcctgccc	ggccggaagg	ggctggggca	gacctccac	103800
tctccatcac	ttccttcttc	ttttcccttg	ctcacagcct	cccgcgcctc	ttttacctct	103860
ccctcttgaa	acttctccct	ctagaacccc	ctagaacccc	agcgggtgtc	ttccctccct	103920
cctcgtgccc	tttcagcctc	ccagccccct	tgcctctgcc	tcccctaacc	aagttagttg	103980
aatgctgtta	ctcgtctcagg	cccacctagg	gaaaatgtca	cacccagcac	ccagaggaca	104040
cacagacagc	acatgagggc	atagggacac	acacactcta	tttgtgcatt	ttgccttgac	104100
cgctgggttg	gcagggaaca	tatttttcct	atttgcctac	cagcttaacc	gtctctccca	104160
gtttcacact	cccagagctg	ccaaaaaat	cccaaccaca	gaatcaggaa	gccaagaacc	104220
aggactgagg	gcttttcaga	aaccatcccc	tggaggactg	cccatatttt	tcaactccaa	104280
aaacccctta	gatgactccc	tgcctcacc	ccgcccccca	ggttctgaaa	gagccttccc	104340
gccagactgc	attgattaac	cattcattgc	ccattttttt	attaatcaaa	gacatatata	104400
attgctcatc	ggagcttgtg	atcagcgtga	ggccttacta	agcagctgac	ttactatcct	104460
tccagcccag	agcacgtgag	ctgacgtctt	cttgggcttg	tgtggccgtt	tccttgccaa	104520
aagctcagtt	tggggagagc	ttcttgcgta	ttagatgcag	tctgcagact	cccaacccca	104580

gctacctgga	tcccctgagg	gccaggaac	tecagetatt	ccaagcccac	tcctcttttt	104640
tttaagagga	agaaatagag	gttacgatag	gggacagcca	gaactgagga	ttttccagct	104700
caccacccaaa	gcacaaaaga	taaaagtctg	caaccaccct	agtgacttga	ctgaatggag	104760
gaaggggtggc	tggggctcctg	taccccaagc	tactcactag	ttatacaacc	tgaggcaagc	104820
tctttggctg	ccccacctgt	aagacgagga	caatagtacc	ttaattatag	gaattgtcat	104880
aaaagaagta	taagatgggt	gtatgaggtc	cctgcatggc	gcaggtgcta	taggcagatt	104940
gtagggtagt	agattttcta	gtctgcagtt	atgtagacag	agccagagaa	gcagctctgg	105000
ggaggaattt	caaaggaact	tgcccacggt	cattctacaa	agctgcagta	ccttcccaac	105060
tctgaaacgt	atgctctcat	caccccgtct	taacaaacat	ttggacatta	gagaaaacaa	105120
gtcttttctt	aaaataacat	tatttatggg	agaaaatcca	caaaaatata	gcattcccagg	105180
acaaacaggg	cttaagatgc	aagattttct	attttactgc	aagacacaaa	gactctgaaa	105240
ttaatgcatg	ccctatcttc	tgctctggca	tacatttttag	tctcctgggg	ggatcagtaa	105300
gtgtggaagt	agcaagggag	aaacagaaaa	aagtcaaagt	aaagagacag	attttagaat	105360
gttaatctgc	aggagcctgc	cagaaagatc	tagctcatgg	gctatctgta	catccaggac	105420
tgaagcacgg	gacacggggc	aggctcgtcca	gggttctgtc	caccttatct	tgttacctct	105480
cttgactctt	agagcctcca	ctccacatct	cccatcaatg	tctgcagaag	acgtggcctc	105540
cactaacaca	agtcttactg	aactgatggg	acaggaat	agaatatcct	ctgaaccatt	105600
cccattgttct	ttggttcgaa	ttccagcagc	tagaaaaggc	agatgctatt	ctgatcactc	105660
tcctgctggg	ctccaatgag	gattaatgag	taacatcaga	gagagaagtg	attataataa	105720
ggtctgacgg	tgcaccgat	gtcttcatcc	ttttctcttc	gcctccttcc	tcattcatctc	105780
acaccttttt	ttttttaatt	gactgattgg	ttcaacaaat	acatgtggta	cctcaggctc	105840
tgtgccaaagt	gccgggatc	gtagagaaga	gattcagtcg	ctgctctcaa	ggggctcatt	105900
ctcttgtggg	agagacagac	aaagaaaacc	aagatttctg	gagtgtggga	atggctctcc	105960
aggcagatgc	tagcacagca	cattgaaagg	cacggaacct	caacaaaaca	ataacattta	106020
ggaaccagct	agagcacagg	gtggtgaaga	aagtggaaag	atttgaggcc	agegtcgcca	106080
tctaagttag	ggcattaaga	attcagccca	catcaatcaa	tcattgtccta	ttgatttcac	106140
cccttaatat	ctctcctatc	tatccgtggc	cactgctcta	tgacagacact	catcatctct	106200
cacagaggca	tcattctgctt	ccaagccatc	gccattctcc	tgcaagagtt	tatttccatg	106260
gttcccactg	gatggcttca	cttaactgct	caaaaccctt	ctgaggtcca	gtcaactggc	106320
tggttaaggac	cagtccaggg	tctggggatg	ccagccatga	gacattgctt	tgaggggaag	106380
agggagcata	gaactggatc	tcctgcatcc	tactgcccac	gtaccaatgc	tggaggtggt	106440
tttccctccc	atcatcagca	agtctggata	tccaggatcc	accctatgga	tgtttttatg	106500
gacagagtgg	gaagatggat	atgttttagt	tagggaaaga	gggtttgcca	aagagggcag	106560
tataagttag	ctgcactcca	tcattccctt	ggcacaaaac	atggctagta	tcctctagtc	106620
ctcaagagca	ccaccttcca	atgcagtccc	tgctgttcca	cagacctctc	tcctcaaact	106680
tcctctgaac	aacctcagcg	agggcaattg	ccactctctg	ggcagagtcc	agatattctc	106740
tcctaccctc	tgacatcact	ttctaaattt	gtatatgtag	ataaactctg	agccattcac	106800
ataaagggct	ttgatttcgg	atacgccaaa	cacataaaca	aacaaacata	agctttcctt	106860
tcacaatggg	ctcatgtaaa	ttaaaatggt	tggttttcca	cctacgggtct	tgaaaggggt	106920
ttctacagcc	tggttttgaa	gtcagaaagc	aaaaggtaaa	tgcaaacatc	atttcacctg	106980
cagagaaaat	tctaatactc	ttgaggcagt	gccaaaaata	atacaagcac	actgctatcg	107040
agccaattac	tggtatctct	gagcttccgt	ctcctcatct	ataaaaattgg	aatcgagctg	107100
tatggattaa	agataatgta	tgaaaactgc	ccaattagta	cactattaat	aaatagcagc	107160
tactgttggt	aacaaatatt	attgacttac	tggaatacaa	agaggaaata	aagtcacatt	107220
tagggagaat	ttcaaagtgt	tcctaacctt	aaaagaaat	aaattagggg	gaaaaccact	107280
aagtaatggg	tgagctcagt	ttaccttgct	taagaagtc	caccctagag	aactgatctc	107340
tagatgacac	ccaaatgcac	tcagtacaac	cccccaagac	tgtctgggct	taaggcaggg	107400
gcttggtattg	tcctgtaagc	tgtgggaagt	ctgttcatga	gccacagtag	acaggaaggg	107460
gatggagtct	tagagctggc	tcttcagggt	atctcctagt	gtgttcaaag	cagttctcag	107520
gagggtaggg	aactactaca	tagccaagta	aatatgaggc	ctccttgctc	tggggagacc	107580
tttctcttta	acagaggtga	atctgaaagg	atacccaaa	aggcactgga	gggtgggggc	107640
cactctggcc	cctcagagca	gccagctcag	cttcagtgga	tgctagaggc	agcagaggat	107700
cagcctggat	cagcctccct	ttcaccatgc	agaaaacaga	gctccccac	cagaccagat	107760
actggaagca	ctgggcccagg	cctaagagaa	agcagagccc	caagccccac	cacaccaggg	107820
cttatgaggc	tactgctacc	caccctccca	agccccagct	ccacttctat	gttcatcaag	107880
caactgttta	ctggtaactg	cacttcccga	tagactttgc	tagaaaaggaa	tgctcagtg	107940
cactgacaat	atctaaacct	gcaactctaa	ggactagggt	ggggaacact	gtgtcaacat	108000
ggaggcacgt	cctacccctg	agaagaaaaa	taaggaatat	tgctgggcac	ctgaatctca	108060
tgagcactga	ctatgtatct	gattcgaagc	gctttttgct	taatctgtca	ctgaatctca	108120
cgtaggtgt	tggtattagc	atctattatc	tgggccaact	gaggcctaga	gggacgaagc	108180
aaactccccca	acatcaccag	gtagcagtg	caggactggg	atagaaacct	gatgctctga	108240
ctgaaactaa	tgcttttttt	tttttttttt	tttttttttt	tgagacagca	tctcactctg	108300
tcaccaaggc	tggagtga	tggtgtgatc	tcagttcact	gcagcctcca	cttcccagg	108360
tcaagtga	ctcctgcctc	agcctcccga	gtagctggga	ttacaggcgt	gcaccaccat	108420

gcccagctaa	tttttttgc	ttttttgtg	agatgggggt	tcgccatgtt	ggccagactg	108480
gtctcaaact	cctgggctca	agtgttctgc	ctgccttggc	cccacaaagt	gctaggatta	108540
caggcgtgag	ccaccatgcc	cagccagctg	atgctcttaa	tctgtgccct	accagcctt	108600
cctgggagag	ttcccaagag	ctacacagag	catgagttct	ggaatcgggt	tgatgggggt	108660
accagttatt	actaatagga	atgaagatgg	gttaattctt	cagacagcac	ccttgattaa	108720
aacaagagag	tagtgctgcc	tctctgtgat	tctgtgtctc	cctgccctgc	tcacacagac	108780
accacaccca	cccacacgca	tgatcatgaa	aagaggaaat	ggatccagga	gaaggagacg	108840
actcctgagt	gaaaacaacg	gggtttttca	cattgagagc	tttgcccaac	accccaaaga	108900
tgaaaagagc	aggaaactgc	tggggccgat	tgaacactgg	acttttgttg	tggaaaaagg	108960
caaagggag	ccggaagaga	ctggaacagt	ttccatgggt	ctggaggatg	gggaagtggg	109020
tagggattag	ctggaggagg	aaggagaagc	tgggggtgga	ggggaacctc	cacttgccag	109080
gagagcacat	gtaggatggg	aaccccagat	gatactcaag	gcatggcatt	agaccagaag	109140
caagtctgtg	gtgaaattag	ggaaggctcc	actgcgagct	gtagacagag	cactggacaa	109200
ggaagtggga	gacccagggt	ccagtccctg	ctctggagcc	ccctgggtgg	gctgccccgg	109260
gcacctttct	ctcttggggc	ctccattcct	acctctgtga	agcgagtgtc	gaacctctct	109320
tagccctgac	ttgctgaaat	gctgggactc	tgtacagagg	ctgacattaa	caggggatct	109380
gtcgtggggg	gctgcaatgt	tcctccagat	gctgcacggg	agagggcaga	aaaggcctat	109440
atggtgagtc	cgccctggga	gcctctgctt	ggaagctgaa	gtggcctgag	agtgactcag	109500
aaaccacgga	agttcccggg	gctgatgggt	tcttatagat	tgtacatgca	gctctcctcg	109560
tgggctgcaa	aaccgcaaga	tgggctgtga	ccactctcaa	ggaaagagcc	ctatctgcaa	109620
aaagcattct	gccctccagg	tcttaaagca	aacacagact	caatccctat	tccttttaag	109680
acaaaattgc	ctcaggggca	tcagggaggg	agcaggcctc	aaatgtgtgc	ctttctagaa	109740
ttctcaatga	aagcaccctt	ttgggtatta	ataatgacaa	cagtaatgac	agtcattttac	109800
tgagtgtgtc	ttttgggaca	ggcattgggc	taagagctat	atgtaataata	tattattatt	109860
tgatgcccac	agccacccca	taaggaggcg	gaggtactat	cattatgcca	acttttaaaga	109920
tgaagaaact	gaggcctcaa	gagatgaagt	aacttggcca	agtcactcag	ccagtaaatg	109980
ggaagagata	gacttcccag	tatccagagc	ccatgttttc	accattatgc	tgaagtacct	110040
cttttctgtg	gccaatgtga	tctgctccca	ggaatcctgt	cttgatgttc	ccttccccat	110100
acagaagtcc	tctctgtgtc	ctcttcagcc	tgatagtata	tcttttcata	ccattctttg	110160
gacatctctg	ttatactact	ccaatgggtg	tccttcccc	acccctccct	gggagcttag	110220
ttgttgtgat	taagtatagg	ggaaatgacc	cacactataa	aaactcataa	gagactgatt	110280
gataaaacct	aaatgcaatt	tattaattaa	cactagaaaa	tgaaccacc	cagcagatgg	110340
gaatcctaag	gctgactggg	cagcacaatc	tctttcagga	aggacaggct	tttgggaaag	110400
gaaatcaata	ccagaagggt	ctttgttgag	tacaaagtca	gaggggaagg	agttgatgga	110460
ttgacacata	ggtgaagctt	gacatacctc	tataaagcct	ccatcctgcc	aaggatcaga	110520
atatccaagg	caggagacca	tctgggtgtc	ctctcctttg	gacagtgtcg	gatttttctg	110580
gatcctatga	agatccttacc	tttctggctg	tcttctcat	gattgtggaa	ggctttttgt	110640
ttccttgttt	gcttagatta	atthctgcgt	atthaataga	actgaaaggc	aatttcccat	110700
tgagaccac	tgaagaggaa	taatcaatac	atactagtgt	tgttgccctt	tgcagagaat	110760
tcacttctgt	gttgtcactg	tatcctcatg	cttcttata	atggagggac	agagatggta	110820
aaaactgga	cttggaagcc	agaccgtctg	ggtttgatac	ctggctctgt	tacttataag	110880
ctctgcaacc	tcgggcagat	tacctaagtc	agtttccct	tctctgaatt	ggggatataa	110940
tagcacccac	ctcaacatct	gtcaagagga	ttcaatgagg	gaatacacat	aaagtgtca	111000
gaacagtgtc	tgccatctgg	taagcagctn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	111120
nnnnnnnnnn	taatttttgt	atthtttggt	gagacaggg	tttgccatgt	tggccaagct	111180
ggtctcaaac	tcccgcctc	aggtgatcca	ccgcctcgg	cctcccaag	tgctgggatt	111240
acaggcatga	gccaccacgc	caggccccc	acaccttta	aacaaccaga	tttcattcat	111300
cgggaagtgc	ctgtggggct	ggtgtggaca	tgtgggtgaa	ggtggcactg	ggagaagtta	111360
ggattctcca	tgacctctgt	tactcatatt	cccacactcc	tcaaattagc	ctgagtctcg	111420
aggacagtct	gatggctggg	caaaccctgc	ggcaaaccat	tccccagccc	tgccctctca	111480
accagagtcc	ttccgatata	tgattctggg	cagctgttgt	taccgtgtc	ctccatgttc	111540
ttccagagat	atccatgcat	gcatcggcat	atgtgtataa	ttattatata	tattatttat	111600
cccacaagct	tttgacatca	atagtagcat	attattaaat	tgttctgtac	attattttat	111660
taacttggtg	tctctgttac	tgtcacaata	aagaacatat	agacctggcc	aggcacagtg	111720
gtcacatctc	gtaatcccag	catttttggg	ggctgagatg	ggtggatcac	ttgaggtcag	111780
gagttcgaaa	ccagcctggc	catcatgggt	aaacccccat	ttctactaaa	aatacaaaaa	111840
ttagccaggt	gtggtggcag	gcgcctgtaa	tcccagctac	ttgggaggct	gaggcaggag	111900
aattgcttga	acctgggagg	cagatgttgc	agtaagccga	gatcacgcca	ctgactcca	111960
gggtaggcaa	caaagcgaga	ctctgtctca	ataaaaaaaa	aaaaggaatg	tatagacctt	112020
ctttattctt	tttgatggct	gtaggtggat	gttctaaaaa	ttgtgtaacc	aatctcctat	112080
tgataattat	taagttatgt	cttcagcatc	atatgaaact	tacaaacaag	gttgcatgtg	112140
ctatccatct	gtaaatgtct	ttttgaacat	ttctgaaata	attgcaggat	aaactcctaa	112200
aatgagaatt	tctgggtcaa	agaggatatg	cattttacat	ttaatagata	tttgtcaaat	112260

tgtcttccaa	agtggctcgt	ccaattaaca	ccccgacctg	taatgaatga	gagtgccttt	112320
tttccccaca	ccctggagag	atgaaaaatt	tatgggccca	ctttggagtg	catgggtggag	112380
gaagctgttg	gccgttatat	aaccctcgtc	attaataagc	ctgggggtgg	ggggggagaa	112440
agagaggtta	gttagtgggt	gcaaacatag	aattagatag	aagtaataag	ttctaattgtt	112500
cgatagcaga	ggagggtgac	tatagttaac	aacaatgtat	tgtatatattc	aaaatagcta	112560
gaatggagga	cttaaaatat	tccaacacat	agaaataata	aatgcttgtc	tgcggccatg	112620
ccaccctgaa	tgtgccagat	cttgtttgtt	cttgggaagct	aagcagggtt	gaacctgggt	112680
agtattttgga	tgggagaaat	gataaatgct	tgagggtgata	gatatcctaa	ataccctgtc	112740
gaacattata	cattctatat	atgtaacaaa	atatcacacg	tatcccataa	atatgtacaa	112800
atataatgta	tcagtaaaga	gagggctggg	cacgggtggct	cacatctgta	atcccagcaa	112860
tttggcaggc	caggggtggga	ggatagcttg	aggccaggag	ttcaagatca	gcctgggcaa	112920
catagcgaga	ctctgtctcc	acaaaaaata	aaaataaaaa	cgaattagcc	aggcgtgggtg	112980
atgcatgctt	atagtcccag	ctacttgga	ggctgatgca	ggaggattgc	ttgagcccag	113040
gagtttgag	ctgcagttag	cctacgactg	caccactgca	ctctccagcc	taggcaacag	113100
aggaagacca	tgtttctaaa	agaaaataat	taaaataaat	aaaataaaat	aaaaagactg	113160
aaaagcagag	tggtaagaga	aaggactttg	gggctcaaca	gtactagcct	tgaaccctgg	113220
ctgttactta	cccatcgtgt	gataagcaaa	tgccttaacc	cctgtgtgcc	tcactttctt	113280
aacatataaa	atagaagtaa	aaatcatacc	cacttcaagg	gtcattataa	aaagccaata	113340
gagataatgt	atataaagct	tctggaataa	tgcctggcac	acagtaggag	tttaataaact	113400
ggaaattcat	tgtttagtctg	ggcagccttc	tgaattctgtg	tcctctttgt	ccactaatgg	113460
ctttgatctg	gatttggctc	aggcaagacc	tggggaaggg	cagagactga	gggcaactgg	113520
aggtataggg	tggctctgagc	ttcccagca	gagttaggct	gggaaaggct	tgggagacag	113580
accaggcagg	tgctgataag	accggaatgg	gaggctggag	cataaggcag	ttcagttttt	113640
cccaaagggg	ggtgtacaaa	acgatctcgt	atgactcctt	tatactgtta	atgttttcat	113700
tttatcgcg	actgaaaaac	aaaaccaaca	tatttaaatga	atgattccaa	ggggattctt	113760
gcttttacaa	aaaatgctaa	agtaggcatt	cacatgttta	aaaattgagt	tgattttaaat	113820
tttaaaatta	ctaagtcata	gtacataatg	tgtgagccac	agctatcccc	aaaatcatga	113880
tagcgataca	ttaatgactg	aagttcttta	aacatcaaca	tacaatgcca	attccagaat	113940
tcagctcaaa	ttctgcaatt	acacaggctg	gggttgaaac	ccagcttttt	tgctaactgt	114000
gtaaaattag	gcaggagggc	taacctcgct	gaatctcaga	tgtctagtct	gtaaaattgaa	114060
gataatgttt	gtttttatct	cacagagtgt	tctggaagat	tcaataaaat	cacagacctg	114120
gaggatgatc	tggctgtgac	acctgtcacc	ccactgatct	ccagagttga	ttcggtgat	114180
caggctggct	gggcagggtgt	cccccttctc	cctcaccact	ccgcatgcat	tcctcccgaa	114240
actgcacact	tggctcaaaga	ggaagacctt	tcctgataga	ggaggaccat	tcctcagtca	114300
agggatatatg	agcacctgtt	ctgtcctgcc	agaatctccg	aaggagctct	cagtaaaatc	114360
acaagatttt	attgtgcatg	gtagcatgag	cctgtaatcc	cagctactca	ggaggccgag	114420
ctaggaggac	tgcttgagcc	caggagtgtg	agaccagcct	gcgcaacata	gtgagacctg	114480
gtctcaaaaa	aaagaaagaa	agaaagaaaa	gaataataat	agtaataaat	cacctgtgca	114540
acgtgctcac	ttctctcttt	ggaatgtagt	aagtgtacct	aataaatgtg	atcattgtaa	114600
tcatacagat	gagcacaggc	taaagcatct	tgactttatt	ctataagcaa	taaaagagga	114660
tttgttttta	cagaactcat	tatgttgtga	aaataatttt	ccaacattaa	caaagaacat	114720
tcttcaagta	aaaggaaaac	cacctatcat	tctcccaacc	ttcaataatt	ttcaattttg	114780
catattctcc	agactttgtc	aacatgaata	cttactttac	atggtcgcga	tcagtgttca	114840
tgcaaattct	tttatcctga	cttttataaa	caaataatgat	gttataaacc	ggtctccatg	114900
tttctgcata	ttctttataa	ttatcatttt	gtggctgcat	aatattgcat	tgactatggt	114960
aactgcagtt	ttcttaacca	tttactgttc	tggggaaatg	gaggataatg	ccagggtcat	115020
gcctggagct	tttttttgte	tattgcatta	tattcttaag	atcaaatccc	agcagtgaga	115080
ttagtcatgc	aaaaagtaat	aatattttca	aggctcttgt	tatatttttac	tagattgttt	115140
tccagagttt	tgcacactgc	tcccagagat	gtaggaacac	agacgtcatc	caaccttgcc	115200
agtgtctgggt	gatgggtgtt	ataaacttct	gctaatttaa	taagtatgaa	atgctatcct	115260
cacacggctt	tcattttctat	ctctttgatc	attaacaggt	tgaactattt	tccaagtatt	115320
tgtttactct	ctgcataccc	tcttgggtga	agtagtcac	cacttccttc	acctgtttat	115380
ctgttgaaagt	cttgaggctt	gttttataaa	tgtgagcgag	cacttcagag	tcaatagaca	115440
ttaattgctt	ccagccagat	ttggccactg	aggctcctga	gcaggggaat	gcatgatcaa	115500
aactacaccc	tggacagatt	aaattaattg	gagaaaatgg	gctgagaggc	agagatatgt	115560
gtcactggcc	tactgtgttt	gatacctatg	tgggggcctg	aactggggca	acggcctgag	115620
tccccacta	ccagttagcag	gaggctccat	gtgtccccca	tattagagct	tgcggcactt	115680
ccatttgccc	cacctcctac	aataccccac	atacatgtac	tcactctccc	ttgcaaatct	115740
agtggcttca	acccacagaa	tttaagggga	aaggaattgt	tctgtcctgt	tcacttactg	115800
cagaaatgag	aaaagcgttg	ttcacatggg	atcacctaata	gaagggatgc	catccccaac	115860
ggtgcctata	aggaaatggg	ggaggggttg	agagttgtgc	aaaatgcaac	agggaatcat	115920
cagagtctct	tgccccatgc	tagaggggtc	tcaaatgaag	agagtctaac	gcaactaatc	115980
tcacggccac	tctaggcagg	gcttcccaat	gcttcccaac	ccccacctcc	atcctagact	116040
ttaccactc	tgctgaacac	agatgttacc	catagcacct	tgcaccatga	ttgtttgatt	116100

agcacctccc	acagtagact	gtgtttctga	taggtcagca	acatttgctg	agcacctact	116160
ctgcaggget	gtgccagggtg	cacaaaataa	acaaagccaa	agacaacatg	gaccttgaac	116220
tcagcaagtt	cagagtcaag	tgggataggg	aggctctctt	cactggaagg	taactccaag	116280
aaacaatggg	actcaacttt	ctaaccaaga	gaactccagg	gagctaaaat	tctgacttct	116340
ggttaagact	ggtgtggagc	ttcattaaag	aagaaaagat	tcacccagac	ttgagttcat	116400
agcctggctt	tgcagctttt	aagtcattga	acctttgatg	aagttatgtg	acctctccac	116460
ccagctgccc	ctaacacott	gcaggggcag	ggctggagtg	caaagggagg	caactggtacc	116520
acagcctggg	aggcaccacc	ccactagtgc	aagccgggca	acctctgccc	ccaaggcatc	116580
cctagcctcc	caactgcaag	catcaatctt	gcacttgga	aggaaacctca	cctttgaaat	116640
ctaggttcaa	atttagaatg	atccagctcc	ttgaagttct	atacagaaat	acagccagca	116700
gccaggcccg	gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gtgggtggat	116760
cacttgagga	cagaagttcg	agaccagcct	gaccaacatg	gtgaaacccc	gtttctacta	116820
aaaatacaaa	attagccagg	catggtggta	cacgcctgta	atcccagcta	cttgggaggc	116880
tgaggcagga	gaattgcttg	aaccagaggag	gcagaggttg	aagttagcca	agatcgtgcc	116940
attgcactcc	agcctgggca	acaagagcga	aactccatct	caaaaaaaa	caaaacacac	117000
gcagcttccc	tccacttccc	aaccacagct	ccatctcaga	caacaagggg	cctcatgtcc	117060
atgcacatga	atatccaacc	aacatgtcta	aggcccaacc	acacccctctc	caaacatctg	117120
ccccttggcc	accctttggc	catgggttca	tgcactggca	gaaaggtagt	tcagagaaga	117180
agcccacaaa	gggcccggaa	gtccacttgg	gctttttgag	attccagggt	ccaggataac	117240
ctaagtgtgg	tctagaagag	agatgcagct	tctgggaggc	acattccttg	gtccttagga	117300
cttcttgccc	ccatggaggg	aaactggcta	gatgagggcc	aaagcagagc	cctctaaagc	117360
acagggctca	gggaaggact	ctttttgacc	agatctaaga	gcagcactac	ctctctgagc	117420
ctgtttctcc	atctgtaaga	aggggacatt	aatagactct	ccccgctaga	gttactctac	117480
atcagccagc	acacgtaagt	tcatgacatg	aagcaagggc	ttaatatata	cccgttgtac	117540
tataaataat	aggccaggcg	tgggtggctca	cacttgtaac	cccagcactt	tgggaggccg	117600
aggaggatgg	atcacaaggt	caggagtttg	agatcagcct	ggtgaaacct	catctctacc	117660
aaaaatacaa	aaattggccg	ggtgtggtgg	cgtgcacctg	tagtcccagc	tacttgggag	117720
gctgaggcag	gagaattgct	tgaacccggg	aggcagagti	tgcaagtgagc	caagatctca	117780
ccattgcat	ccagcctggg	tgacagagca	agactgcate	tcaaaaataa	taaatacaca	117840
catacatata	tacatacata	catacatata	tacatacata	catacaatac	atggacaggg	117900
accctaaaaa	tgagacaggg	aaagagaaaa	acatgttctg	acaaccttgc	cctttatact	117960
aattttaggtt	ttcttgccctg	ttttagaaag	ggcctggaca	ggagccctgt	tcccctcagg	118020
ccaggcagaa	caaggtgtgg	aactcactgt	ggaaggggtc	tgggtgacaa	gtgcagcccc	118080
gtccctccac	ctcccagcac	agtaggcagc	acgtgtctcc	attgactggc	tcaggagcag	118140
gcctggtgac	cagtgggaga	gctgaggagc	ccagggtggg	gtctgaagga	atccctagaa	118200
aatctgattt	tccccagggg	cccacatcac	gtgcccagag	ctgggaaagt	ggaggcagca	118260
tggtgatctag	ctgagaggct	ccatttttgg	tagcttctag	tttgggagtc	acagagacac	118320
ctggatgata	cgaagatgta	gctttgcaag	actctctaga	acatggagtc	caagatatte	118380
ccttcaatga	tgggacactg	aagcccacag	aggagaggtc	tgtcccagtt	actcagccat	118440
tcggaggcag	agaccaggct	agaactcagg	acttttaatt	tggaccagga	ttccttttat	118500
cacagtgggc	agccctagca	agtgccagggt	agggtggaac	tgtgaaggtc	atccgagggg	118560
tagtacacgt	gggtaggaag	tcatatctaa	gaactgaccc	ccagacctgg	ctctgccact	118620
cactccttat	gagaccacag	gtgctgggtg	cagtgggttca	cacctgcaat	cccagcactt	118680
tgggaggcca	aggcaggcag	attgcttgat	tccaggagtt	cgagaccagc	ctgggaaaca	118740
tagtgagacc	cccacctcta	ccaaaattag	ccaggcgtgg	tgggtgtctg	ctgtagtccc	118800
agctacttgg	gaggctgagg	tgggaggact	gcttgagcct	gggaggcgga	ggttgccgtg	118860
agccaggatc	atgccactgc	acaccagcct	ggatgacaga	gtgagacaga	atgacacact	118920
gtctcaaaat	aaataaataa	atgacagcag	atcatcattt	ttctttctgc	ctctagactg	118980
caatgcctat	ttctccaggt	agtcactagg	ataaaaagtaa	aaataatatt	atcagcattt	119040
accaaataca	gggtcagcta	ctctgttatg	ttctttcatg	ctttgtttct	tttaagcctc	119100
aaacaactct	atgagctggg	aacaagtatc	gtccttcttc	ctcccatctt	atttatttat	119160
ttatctatgt	atttatctat	ctattcattt	atttatttat	tttgagacaa	ggtccttcta	119220
agtcaccagg	gatggcctca	aacttgagct	aggaactagt	gtcaccaccc	cccaatttct	119280
tttattgatt	gattgattga	ttgattgact	ggttaatttt	gaggcagggg	tctcgctaag	119340
tcgacagggc	tggctcttgaa	ctcctagtct	taagcaatcc	gcccgcctca	gcctcccaaa	119400
ttgctgggat	taccaacacg	agccaccatg	cccagcccct	cccattctct	gaataggaaa	119460
actgggggtt	gaaaaggtaa	gcgacttgcc	caaggtcccc	tccttagcta	gagagcttca	119520
gagccagggc	acaaacccat	caaagcctgt	gctctcgccc	attgagccac	cggacctcgt	119580
acactaaccc	ccaagtgttc	tacacagtga	aggtgacaaa	gaggtgaagg	gaagagccag	119640
ggagggttctg	tggactcact	cgggtgggtat	gccagagggg	aagggggatc	ttgggtggca	119700
cattgagagt	agctgcgctg	ttagtaagtg	agaactcgga	agtccagact	catccagctc	119760
gtgccaataa	acccctcctt	ctacctggtc	tcctttccaa	agccagctgt	tctccagaca	119820
atgggggtgg	cggggggcgg	tgtcctcctc	cttctcaggg	aaaatccgac	gctgagccca	119880
tctccagaga	tcttggtctc	ccgtgggggt	gcagatccac	ctagagccac	cagagggcgg	119940

gccagcactg	cggccaaggg	ttgaagaccc	agcacaccaa	agcccggcca	agcctccage	120000
ccagtgtcca	agagtccagc	cagaggccga	gtcctcgatc	tcaaaatgtc	taactgcaga	120060
agcccaactc	atgttcaggc	atgatgtgtc	tgattctact	gggacaatca	ttgccaccaa	120120
agaattactg	ccaaaatagt	aacgacatta	gctacctacc	acccctccac	acaccaacac	120180
acctcatttt	accaagcact	ttctcatgcc	tgggggtgct	ggagacttaa	tgagcctcg	120240
cgatgaccgg	gtagcctcac	cgtacagatg	aggaaactga	ggcacaagga	aggggagtag	120300
attgtctagg	gtcacctgga	gaactctgat	ctccagactc	aaatttccaa	ttcgtccccc	120360
ctccccccaa	ccctaaccgg	agctaggtgg	ggtggggaca	gcaaattgtg	atggggggag	120420
gtaagagggg	tcagagtgtc	ctacagagaa	gaccaaatgc	attgtggcac	ctactgtaaa	120480
atgagaccag	ccagccaccc	acccaccagc	cagccaccta	aaagtcttca	gtgggcacct	120540
gctggaaaca	cgacaatgga	tgacacgagt	tccctgccct	caaaaagctg	gtagtctagt	120600
tgggggtgga	gggggtgagt	cagcagataa	ttatgggaaa	ccgtgacacc	tgtataaggg	120660
gcggggatga	gcagaggggc	tgcgactgcc	tggagccagg	gattcccggg	cggggcttcc	120720
ctttcctcgc	agctcgtccc	aggaggagga	gctccccccc	agcttcgggg	ttccgcctgc	120780
cttggggggc	cggggtcccc	tcccacccct	ccccgaagag	cgcgggcccc	gggaaccgat	120840
gacagcacac	ctgagtcagc	ccgcgcgccc	accgcctctc	agcgtctgtc	tccgcctctt	120900
gtgatatttc	gctccccggg	agccagccccc	cctgcgctcc	ggaggcagct	cggcaaacaa	120960
acccagcgac	agattgtgcc	gcggtcattt	ccggggaagg	acgccaacc	ccaccctgct	121020
accccccaaca	ctccctcccc	gccgcgcgct	ccaggccctc	cccccaggcg	caggccctag	121080
tcgggggtggg	tcctggggaa	acgcagggtc	ctgtcctgcc	tcctggaaat	agggggagcc	121140
ctgggtgagg	gaagacggga	gccccagaga	cttttcttcc	tgattccgaaa	ccgagaaggt	121200
acgagagggg	cgggaaagga	aatttagggg	cacagagagg	agctgggggg	cgggagcctg	121260
ccgaaaatgg	aaccagcagg	gggcacccga	gagccgaggt	gcccacgggc	cgggagcctg	121320
ggaaatgaaac	tggggaagag	ggggagagaa	agggaggcag	agacaccgag	acacacagag	121380
acgagagaca	gagacgcagg	gagccccgcg	gggaggagga	gagagacgaa	gacacagaga	121440
gacagtgaga	aagacagaag	accgggcagg	gaaacagacg	agtagagaca	gaaaaggtcc	121500
gagagagagt	gagggagggg	gggaacagag	agacagagac	cacgaaatat	gagtaagagt	121560
cgggggagaa	aaccagagaa	atcgaatgag	aacgcgagaa	gaacgagaga	ccgtggaggg	121620
agcagagaat	gaatgggaag	aataagacca	acatttatca	agagccgact	gtatgccagg	121680
cactgcattg	gaccctggca	cggataggaa	aggaggagcc	gcggcgcggg	cagcgggggc	121740
aggggcttct	gtgctcgcgg	gagcggcagc	ccaggggggt	cagcagcccc	ggcacccggc	121800
cacctgcggc	tcagcagcc	ccaaccccg	cagcgctgcc	tgcccaccgt	accggaagcg	121860
gtcccccgga	gggccccgag	cctatcctac	gcggggcggg	ctccgcggac	gcgcggggcc	121920
gagtcaagac	ctgggtcaac	cgccctgcag	cctttgtagg	gaagtgccta	ggtgatgggt	121980
ctgctgatac	cgctgtgac	caggccatga	agggccagag	gggctccagt	gagaccataa	122040
tcgcgccctc	tttaaaaggg	ggtagaggaa	gttcacgcga	agccaacagt	cttctcccca	122100
gctttgggtc	ctctcctgca	ccccgcgggg	agataaggtc	tcccctcccg	gacacatcat	122160
acatacacaa	aaaaacgcac	acactcgcac	gcgcgcccat	ctcgcacccg	cttgtaaagt	122220
cactaagggg	catacacaca	ccgggcacat	atttctttcc	acccatcccc	aagatcgcaa	122280
gcgcaaaacc	tcgcacagcc	tcacgtttcc	caccagctca	gacatgcacg	ctggcggaact	122340
ttcagcggct	caccctgtgt	cacactcagc	tgcccccccc	cccgttccc	caagcccgtg	122400
caaagggtaa	cgggcaagca	tcctgagtca	cacctgcaca	agcatccttg	cgcgacggtg	122460
cacgctcata	tgactctgat	cttgacgcga	caaactcttg	catatactat	tcttatagtc	122520
gcacactggg	cttgaggtct	gggagtgga	ggaaaagtgg	aatcttgagg	ctgtcccagg	122580
ggacagaaat	gctggaggct	gggacactgg	cgcgagggac	gcggctgggg	gcgggggagg	122640
gggtgaccca	gaagctcatc	ttctcctgga	aagttgggag	gggggaacag	gacaagtcca	122700
cggcgttctc	ctaaactacc	gcattccccc	aagaagggat	ttctctagaa	gagtggcgcc	122760
gcgaggacga	tcgaacacag	tcctccgggt	cgcttaagcg	ggggggaggg	gggcgggggtg	122820
gaggggggta	gaaagccgct	cccgcctcct	agtggtcgag	aaagggttaa	gtcggcaagc	122880
cagcaaacga	gggaggagcc	agcgagtgcg	ggaaggagtg	ggggtggttg	ggaagagctt	122940
cctcgctgtc	cccactctcc	ctcggctagc	agcctgggca	cacggacaga	cggactgacg	123000
gactctcgag	cggaacagcg	agctagcggg	gcgcggggcg	tgggcgtcga	cgccagcccc	123060
cagccttccc	cgccccgtcg	cgccccggcc	cgccccggcc	gggcccagtg	ctcctcccga	123120
ggcccgagc	ccgggcggcg	cagggtagag	cgcccgggcc	cgccacgca	gcccggggac	123180
tcccgggccc	tcccggagcc	ccgcggggtc	cccgcgtgc	atccggcggg	ctcagggagc	123240
gagtgggagc	gccctcccc	cgctgcccc	tcccccgagc	atcgagacaa	gatgctgccc	123300
gggctcaggc	gcctgctgca	aggtaaaga	gccagcggcg	ggagagcgga	gggcatcctg	123360
gggagagaag	cagggcgtcc	cctctttcag	ggattgaggg	tggggcagtt	ggggaggtgg	123420
ggtaacctgg	ggaaggggaa	aagctcagcg	ctggggccgc	gccccgcgcg	ccagggtgtg	123480
tctcagcagg	agggcacttg	gctgggagcc	cgcgggcgcg	tgcgaggagc	tcgtgaccga	123540
ggtgggacgc	agggggcagg	tggacccggc	ccggagcggg	gagggaggct	caggttccgc	123600
tgtccccgct	ccacctgctc	cgggggacgc	tgaggactcg	ggccggctgg	ggaagcgccg	123660
actcagcaac	tcctcctgcc	cggtgcctca	gcactttctg	gccacctggg	aagacaggag	123720
atgtgggtag	ggggctgtct	ggggaggtag	gaggcgacga	gggaaatcca	agtggccctc	123780

tctggttagga	gagatggagg	gcgctagaaa	gaggatagtt	ctactgattg	agtgcacagat	123840
aagggtgtgg	gccagagact	gggggtgggg	tggggagggg	tcagggggag	agggatagga	123900
aggagaactc	aaagatggag	aaagtgtgga	gggaagctca	aaggaggagg	gagatggagc	123960
gggggagggg	gagaaggaat	aaagggttaga	tgggaaaagc	gtggagggaa	gtgggaccca	124020
ggtgaagacc	aaggaagagg	gaaggagagg	aaagaccaga	tcaggggagg	gatgggaaga	124080
agactatgga	cagggaccca	gaatcctggg	atggaggtag	cgggaaagag	aatcaggact	124140
gggacctgg	ggactggaat	ggaaaaggag	aatggaaaaga	tcagaaacca	gagaaggatg	124200
gggatggtag	ctagagaagg	ggtatcagga	accggcgaag	agggttggag	acagggaaacc	124260
atggatggga	gaggggctgg	agaggaggga	agaggaggag	gaagagagag	gctgagagag	124320
agggactggg	gattgggggt	gctgcccagg	gatgagacaa	agaggcttct	ggttaaccact	124380
tccacgtggg	aagccctcca	ttcccaaagc	gcctgcctgc	cacatttctt	ctctcagggg	124440
gtggctgggt	ggccagatgg	gggggtgcttt	gagctcaggg	ccctgggggt	ggctgtgagg	124500
gacagagggt	gaggactttg	gaaggggagt	gacagcctcc	gaggggtgggc	aaacaggctg	124560
gctcctgtgc	tgccatttat	ttatccggcc	cggacgttgg	attctgcagc	cgctgccgcc	124620
accacggtgg	ctgcttattt	tggggtgtta	cattctggca	gagtgagaag	ctgtttgag	124680
cagctctaaa	cctccgtcac	ccgctcaggt	gcctccccag	gccccctgcgt	cactggcattc	124740
accaccacct	ccatcccact	cctcagctcc	cacctcctca	gccccctgcc	cctcagcatc	124800
tgcccgcagg	ccccagccct	tccctgaagc	agcccgttgg	gtgtggagcc	cttgcttctc	124860
gtctgggacc	ctgtgcccct	ccttcagag	cgagaggcct	ctgtgcctt	tccagggagc	124920
atccttttct	gggaccactc	tgcaccagcg	actctgccct	gtgggtgggt	agcctggatc	124980
ctgcccccta	ctttgggtcc	agttttcttc	tcctcaagtt	ccttcttcta	caggggcctc	125040
cgccccaaag	agtggcctgt	gggctgagaa	ctttgtttct	gagccttggg	actccaaggt	125100
ttgatagcca	gagtcctgga	cagtgttccc	tttgtgaaca	gatacttttg	gctctggaca	125160
cttcagcctt	ccgggatcaa	taccatgttc	tggcctctct	tggctccctc	ccctggtcag	125220
ttctggccat	atattctgga	caggggtcat	ctcttcttga	ctcccacatg	taatcactac	125280
tctagaacaa	ccgcaactgg	aagcctagga	ggtgaaagtt	gcagagagag	ctggagtccc	125340
ttccttgccct	tgaccctgaa	tagccaaaca	gactcagcat	tgtggctggc	ccagccctag	125400
gcacctgggt	gcaatttctc	tcctgtcttt	acctcaaggg	cagtgtctca	cacattcagg	125460
cgtgggttct	gcggaggatg	tggccacctc	ttaaagaaag	atcagagtgt	ctctctgaca	125520
tgggcttgat	gtccctcttt	tccaatctgg	gttccacctt	gtactagctg	catgacctga	125580
ggccactgtg	tcatgtttct	ggggctccct	tccttcatct	gcaaattggg	ggccacaata	125640
ttgacctcca	ggggattatg	tgtgttgtgt	tcaatgtata	aagaagttaa	cctgtacaaa	125700
tgcagtgcct	aggacaaaat	aggtgcttct	tgttttctct	ctaccctgct	gtactctccc	125760
ctgcagctct	agccatcccc	tgtgactttt	agaggagggg	gtgagcagag	aggggtgggg	125820
aggctgtac	aaagggtctt	cctctgtcca	tgaagtagtg	gagggatgaa	atgaaggctt	125880
ctgagaaaga	caatgaaggc	gagctgtaga	gacctggtca	ggaggcctgg	ggtgctcaga	125940
aactcacact	tccccctccc	agccctcaat	ggtgttaacct	atgatgtgag	gggtcggctc	126000
taggtggcca	ccgaggtatc	cccccttcca	gctctgatac	tctgtgcata	ttgccccagt	126060
ctccacoggg	aattcacaaa	atgaaggcca	ggagtggagc	cgtggtcctc	gggagagaca	126120
ggaggcctgg	gcctggaggg	aaggagtggg	ggtgctgagg	aggagtgaga	acaggggggtg	126180
gggaagggac	gtggcaagaa	agaaaagggc	acacactggg	cagggcaggg	actgagggcg	126240
ggggagagag	ggaaaggcac	agctctctag	tcccccaacc	ccccagtccc	accacctctg	126300
ccctggagtg	ctcgetccag	ccccagcagg	cctggggcag	tgaagcccag	agccccctcc	126360
cctccccctc	tccttgccct	cagtgcagag	cgctgcgtga	attatggatg	agctccttgg	126420
gttacagctg	ctttgcacgg	cagtggcaag	ggccagaaat	ggcaacagag	tcactgttat	126480
gcagcagctg	ttatggagga	gccccagca	ccgggtcgct	cttcagagag	cctgcaggga	126540
ccactatcat	gggctggggg	aggtgagccc	tgggttgggg	agacatggga	acaagatgga	126600
aggagagtgg	ggaaagagaa	gagaagtgtg	ctaagtgtgg	caggtgggga	gcaggagagt	126660
ctagggagag	aaagaggagt	aggcaccctt	gccagctcct	gcagagttta	ccctcaaggc	126720
cggaaggaaac	cctgatgcca	ggggaatggg	ccttgccctc	gagattgcac	atccttccct	126780
ctgtctctcc	tggggcagcg	gtcagtcagg	aggctggggg	aaagctctgt	aatcctccag	126840
gggctagcgg	ccatcagggc	tcacactctg	gtgagcttgt	ggataagggg	taggattaag	126900
ggatcagaga	aggatttggc	ttcttttggg	gtcaagtcct	tagggaagtg	gagatcagag	126960
ggtgactctg	acaggaaggg	aagtgccctg	gctgggcata	aagagacttt	tctggccctt	127020
tccttgccaa	cactttgctg	tgtgaccttg	ggtaagtgcg	ttgctctctc	tgagctccag	127080
tcatcacctc	agtagaactg	atgcttgaac	cagaggaatc	gaggggacct	ttgcggcttt	127140
gaaatctcca	gttctaagcc	ccaaacctca	accctcatga	aacctactca	gggtccccac	127200
tgtgcttcca	cactccacct	ctgcctgggt	cagatgaggg	gtaagagaca	ttgctcctcc	127260
acccccacgtg	ggtctaagaa	actcgggagg	agaaagtaat	cgtgaaacgc	cgcacggggg	127320
aggggtgaga	agggccgaga	aacgcggagg	tgggtgtgaac	gaatggaaca	gcagccgctg	127380
tgtcactgag	tattacatca	cacccagcct	acacacgcac	ggggcccgcc	gctcacacac	127440
acgcggagga	cagccagcac	gcaccgacgc	agcaccgacg	cagcgccagg	aggggcccgg	127500
gacactcacg	gtggggccca	aaagcgagga	gcagcacact	gggagtgtgg	atcttccacc	127560
ccgcacctgt	gtgctccccc	ctctggagga	ggaacaccag	ggcagctggg	atgccagcgc	127620

cacactcggg	gectgtcagt	cccatgcgtg	cacacctggc	tgagcagcac	tgcatttggg	127680
gagcacctgg	ctcacgccac	tacccaaaat	cacagataca	tacacacatt	cacgcacacg	127740
gcaacctcag	gagcgtgaca	caacacacac	aaaaccacca	ctaagcaagt	gcaatttgcg	127800
gccttgagga	ccccacactc	aaaatcacca	acccctcagt	ctctcccagg	gtctctgaac	127860
cccaaggagc	cccaggatgt	cagagtgcag	aaacaagtct	tcctcccctc	tgccctcaaa	127920
agcctaggac	gttgcttgaa	gcagaagggt	ttcagtcact	gtgtgcccag	ggaatgactg	127980
cctggccttg	ggggtgcagg	ctcccttttt	ccccaggcaa	aactgccaga	agaaaatccc	128040
aggagtccac	tggaaatcat	aagaaagtgt	agaggccaag	ctagttccgg	cctagaactt	128100
tatcagctat	agtgcgycfa	aaggccaggg	atgatgggag	gccctgcacc	cctattaaaa	128160
tatgagtaca	gacacctgca	ctccactctc	tagccccagc	gctctctggg	cctgcttttc	128220
catcagtatc	ataataagga	tggatcatat	ccaaccttca	aaagttactt	tgggggaaaa	128280
aaaaaaaaaa	gctttggctg	gatgcggtag	cttatgcctg	aaatcccaac	actttggggag	128340
gccaagggtg	gaggattggt	tgaggccagg	agtttgagac	cagactgagc	aacatagcaa	128400
gaccccatgc	ctacaatttt	tttttttttt	tttttttttt	tttgatacag	agtctcgctg	128460
tgtcacccag	gctggagtg	agtgggtgcga	tctcggctca	ctgcaagctc	cgccctcccg	128520
gttcacacca	ttatcgtgtc	tcagcctccc	aagttagctg	gactacaggc	gcccgccacc	128580
atgccgggct	aaattttttt	tttgatattt	tagtagagac	ggggtttcac	cgtgttagcc	128640
aggatgctct	cgatctcctg	acctcgtgat	ccaccgcgct	cagactccca	aagtgtctgg	128700
attacaggcg	tgagccaccg	cgcccgccca	aaaattttta	aaaaattagc	tgggtgcagt	128760
ggcacggggc	tgtggtccca	gctcctcagg	aagctgaggg	aggaggattg	cttgagccca	128820
agtgatccaa	gctgcaataa	gctgtgatcg	taccactgca	ctccagcctg	ggcgatggag	128880
caagacctgt	tctccaaaag	aaaaaaaaga	agaagttttt	aagtaactgc	gaatgaggag	128940
agcctggggg	gtaaaaatgca	gattccagg	ctgtccccc	aggaattctg	catagttcct	129000
aggactggct	ggtggcctca	cttagagacc	cgacccttaa	ggccctccc	ggcacaaga	129060
ggctctgact	ctgcaagggc	gaaaagtaca	ggaaagtaag	ggcactgggc	accagtgggc	129120
tggcaagacc	agaccccaga	gtgagtccat	ttcacacggg	cctcagatct	ccaaagggtc	129180
ccaagttact	tccagtcatt	ctccaatggg	gtgactttgc	ccccagggg	acattttgca	129240
atgtctggag	acatttttgg	tgtcacaact	ggaggcaggg	tgtgtctggc	atctagtggg	129300
tagaagacag	agatgctgct	aaatgcctta	tatagggtg	ccccacaac	gaggaactat	129360
ccggcccaac	tgtcaatact	gaggcagaga	aacctgacg	ttagtctttt	gacattaatc	129420
tctagacaag	gtcaaacatg	caatagtga	aacaggaatg	aagagatgat	cattcttcaa	129480
ccaatttgca	gtgctttcta	caatggcctt	ttggcattat	tttttaatat	atgagaagcc	129540
tcagaaagtg	gaagtggcca	ggccacttga	ggctataacg	ttgtcccctg	agcccccaga	129600
catggggagca	ccagggtctc	aggcctttat	ttttattttc	tattttttcc	cctgaaacag	129660
ggtcttggtg	tgttgcccag	gctggagtgc	aagggtgtga	tcgttgctca	ctacagcctc	129720
aaactcctgg	gttcaagcga	tcctcctgcc	tcagcctccc	aagtagttgg	gactacaggc	129780
acatgccacc	atgcctggct	aatttttttt	ttttctttgt	aaagacaggg	atctccctta	129840
tgttgcccag	gatatgtctc	aactcctggc	ctcraagcaat	cctcctgcct	tggcctccca	129900
aagtgtggg	attacagggt	tgagccacca	tattcagccg	ggtctaggcc	ttttaccaag	129960
ttggggggct	ggcccccagc	tggcactcct	gccctggaag	cccacctagt	aagttctgct	130020
tccccctccc	acagctcccg	cctcggcctg	cctcctgctg	atgctcctgg	ccctgcccct	130080
ggcggccccc	agctgcccga	tgtctctgac	ctgctactca	tccccgcca	ccgtgagctg	130140
ccaggccaac	aactttctct	ctgtgcogct	gtccctgcca	cccagcactc	agcgactctt	130200
cctgcagcaac	aacctcatcc	gcacgctggc	gccaggcacc	tttggttcca	acctgctcac	130260
cctgtggctc	ttctccaaca	acctctccac	catctacccg	ggcactttcc	gccacttgca	130320
agccctggag	gagctggacc	tcggtgacaa	ccggcacctg	cgctcgctgg	agccccagac	130380
cttcaggggc	ctggagcggc	tgcagtcgct	gcatttgtac	cgctgccagc	tcagcagcct	130440
gcccggcaac	atcttccgag	gcctggtcag	cctgcagtac	ctctacctcc	aggagaacag	130500
cctgtctccac	ctacagggtga	gcctgccctg	ccccaccct	cagccccctt	ctgggttctc	130560
ctctctgtgg	gccccctctg	tccccgaccc	tggcgtgcgt	ccctcctctc	tccccaggcc	130620
acccttctctg	cctcagcatc	tccattttct	tctgtctatg	tctctttttt	ctcttacatt	130680
ctccaggggc	tttaactttt	cccttctgcc	tctctacctg	tttaggtccc	ttgtgttcc	130740
tctctctctc	tctccctcta	actccacaac	cttcacctct	ctgctctgct	ctgtctgtct	130800
gtctatccct	ttccatccat	cactgcctct	ctcacaaact	tgcctcccc	atctgtcttc	130860
tgcctcttct	gtctgtctcc	cttcacacac	ccactccgca	tacaccccca	tgtctgtctg	130920
cgtgtgtgta	tctgtctctt	tctgtgatct	cacgtgtttg	ccttcagggg	actctgcctt	130980
ccccagggtg	cccctgccc	aaggcctttg	cagctgtttt	tctcaccac	cctcaagtct	131040
gcccacatca	cgggtgaagta	gagagagaag	gcagagccac	agccactggc	atccccacaga	131100
aagttgcgct	tctctccaat	tcactgggca	atgggacggg	agaagccca	accccttcta	131160
gattcccat	ttccaaacct	gtcatctcaa	tgcaggggaa	gaaagaaaag	ggtaaatctc	131220
tgttatgcag	ctggagaatg	gatgctctga	aaatggaagg	aataccagta	attgttatct	131280
attgttatta	ttattgatct	aattattggt	tattgttgtt	atgctgactg	tttgacacgc	131340
aaatcatccc	actccatttc	cccaggaagc	aataacacac	ccctcaaac	accctgagag	131400
aaaatcttcc	cttggtctaca	gagcctccgg	ctggaagggg	gtgaaaatat	ccaaattctg	131460

ccctctccct	acttgaacct	ggaacgtgct	tcctctgect	catccagggc	tagtgcctaa	131520
ctagtattca	atctgctagt	tggaaaatca	ggtcagtgct	gatgatgcta	atgataataa	131580
caatagccat	aacaacctaa	caaacatact	gagcaccac	tacgagctag	atgctaagaa	131640
tacagtagtg	aacagaacag	accaaaccct	ctgcccctac	agagatacca	ttcccatgag	131700
gagggaaa	agtaaaatgc	acggatatatt	ggaaaaatat	gtcttatatt	attcttattg	131760
ttgcctaaat	agtgacagta	atagcagtag	ccgccaccac	ttagtgggta	cacagggtca	131820
gccacagtg	caagcacttt	ataggtatcc	actctgccat	ttacaagcgt	gtgacatttt	131880
ttttttttac	ctcctcagac	ctcagttttc	tcatctgtac	aatggggtag	caagagcacc	131940
catctcctag	ggattttgaa	agcattaaat	gcatgaataa	tttgtaaagc	acttagaata	132000
gtgcttgcca	tacggtaagt	gctatataaa	tgcttgtaa	aatactattt	taaaaaaaga	132060
aacgagcctt	atttaacatt	ggtttcagtg	aagtggccca	acttggaact	catcctgaag	132120
atgtgggtca	acttcaagga	ttataactaag	gtcatgagtg	agtcccagaa	attgcacctc	132180
acagtttatg	aagtgcactc	agccacctca	tctcatttct	acagcccagt	tgggagatta	132240
ttttcacctc	cttggttaaca	atggagaagc	tgaggctggg	ggccctgaag	accctataga	132300
gatatagtca	cctccaatca	taaatctttt	caaccattgt	cggtgtgacc	ggaggcttat	132360
gtcttctcac	catcatgttg	agcctcacia	caacctgggtg	ataggggacag	ttagggggcac	132420
tagggacatg	gaatgaatgt	tcctgaggcc	acacaccacag	gaagagctgg	cgcttgaaac	132480
tcattggctg	gctacaaggg	gacagtactc	tggagtacaa	ttgagcaggc	tcatttttga	132540
aagcacacag	tttggaactca	gcaagacctc	ggttcaaatic	ctggctccta	tatatatgac	132600
tttggaacaaa	ttacttaaac	tctctcagtc	tccatttctc	catctctaaa	atggcaatca	132660
ggatagtaact	taataataat	cttttttttt	tgagacgacg	tcccactcta	tcgcccaggc	132720
ttggagtgcag	tagtgcgatc	tcggctcact	gcaacctctg	cctcccaggc	tcaagtgatt	132780
ttcctgcctc	agcctcctga	gtaactaaga	ttacaggcat	gtgtcactac	accagctat	132840
tttttgtatt	tttagtagag	aagggtttca	ccatgtttgg	caggctggct	ttgaactcct	132900
gacctcaggt	gatccactca	cctcggcctc	ccaaagtgtc	gggattacag	gtgtgagcca	132960
ccatgcccag	ccaataataa	tccttattta	agaagttttg	taaggattaa	aatgtaaggc	133020
atthagcaca	aggattaaaa	tgtaaggcat	taagcacata	tgggcactat	aataataatt	133080
actactacta	ctactactaa	tactgagatc	aaatactact	acaaattgat	catgcattta	133140
atgctttcaa	aatctcctta	tcaatatata	ttagtatttt	aggaggaatt	tggagtcaga	133200
gggctgagc	ttgaatcccc	gatctactat	tttctgactt	atttaacttt	aagcaggttg	133260
ctaaccctct	ctgaacctca	cttactttat	ctgcaaaactg	ggaataatga	aaataatacc	133320
ttccaccaag	aatggctgta	aataggaac	gagttagttg	atagaaagcg	catagttcag	133380
gctgggtggg	tggcccatgt	ctgcaatccc	agcacttcgg	gaggccaagg	tgggtggatc	133440
acttgaagtc	aggagttcga	gaccagcctg	gccaatatgg	tgaaccctg	tctctactaa	133500
aaatacaaaa	attaggcagg	cggggtggca	ggtgtctgta	atcccagcca	ctagggaggc	133560
taaggcagga	gaatcacttg	aaactgggag	gtggagggtg	cagtgaactg	agatcgtgct	133620
actatactcc	agcctgggtg	acagagcaag	actctgtctc	aaaaaaggaa	aaaaaaataa	133680
aaaagccccat	agttcagtg	tgaagaaatc	atgtttattt	gaccccatcc	tccattgact	133740
ctcaggccaa	caacagcaat	caggacctga	ggtcagcaaa	ggcttgggca	gaggggacct	133800
cagggtggaca	ttgggggtctt	ctgaaatggg	aagtgtttgt	tctctacgcc	cctggcatga	133860
atggtaccag	gcatcatggg	aagggaagcaa	cttcacacct	ggccttttat	agaggagatg	133920
gaaaacacag	cctctgcctg	tgaactgcct	ggttagggctg	ggctgggaga	tgccaaggc	133980
agggtaggaa	acatgggctg	gggtgagatc	cgcagggtgc	aggtgtgacc	caagatggag	134040
ccaggcctgc	cccaaagggg	agccttgagg	gaaactccac	cagaggacca	cagcttttca	134100
gaatggggaa	gggccaggca	ctgtgccagg	tgagttcatt	catcaacaga	tatttactga	134160
gtatctacca	catgccaggc	aatgttccag	gtgccaggga	ttcaggagag	aacagaaaca	134220
gtggccctgt	tctcccagag	catattccct	actcaagtgt	agccagatga	taaagacact	134280
tgttttcttt	cttttttttt	tttgagacga	agtctcgctc	tcttgctcag	gctggagtgc	134340
agtggcacga	tctcggtcca	ctgcaacctc	tgccctccag	gttcaagcga	ttctcctgcc	134400
tcagcctccc	aagtagctgg	gattacaggc	atgtgctacc	atgcctggct	aatttttgta	134460
tttttagtag	agacgggggt	tcaccatgtc	ggccaggctg	gtcctgaact	cctgaccaca	134520
ggtgatctgc	ccaccttgge	ctcccaaagt	ggttgattac	aggtgtgagc	caccgcaccc	134580
gccgacactt	gttttctctt	tcagtcatta	cagtgccctg	catggttttt	gtttgttttg	134640
ttttgttttg	ttttgttttg	tgagacagtc	tcactatgtc	acccagctgg	agtgcagtg	134700
cgcaatcttg	gctcaactgca	gcctcaacct	ctgggggtcaa	acaattcccc	catcttagcc	134760
tccccagtag	ctggaactac	agacatgtgc	caccatgtcc	agctaatttt	tctattttat	134820
agagacgggg	tttcacccatg	ttgccaggc	tggtctcaaa	ctcctgaact	taagcaatcc	134880
acccgcctcg	gcctcccaaa	gtgctgggat	tacaggcatg	agccaccgta	cacagctggc	134940
ctgaatgggt	taaaaatagt	ctttatgtct	aagcagatca	gatctcagtt	tgaattccag	135000
ccacacctct	aatttgctct	atggctttgt	gcaagttatt	taaccactct	gagcctcgat	135060
ggacccatct	gtgaaatggg	gataacctgt	accttgccga	gcaggggttg	tgaggattaa	135120
aggagatact	actgagctca	cagcccaatg	tctggtacaa	agtgaagtac	caatgaatgg	135180
tagctatcca	taaaccaccag	ggaggacacc	aactgaagct	cagcaaaata	aaagcacagt	135240
ccaaggtcac	ccagctagta	aggaacatga	cctagaattg	gcccaggctc	gtctgactcc	135300

agagtgcagt	tgttcagagg	tctctggagt	tggaagccac	gttccactgc	atattagctg	135360
ttggacccta	ggcgagtcac	ttcactttctc	tgaggctcca	tctcgtaatc	tctgaaatgg	135420
agataataat	agtatccacc	tcatagggtt	gtgacaatta	agttactata	taggatctgt	135480
gtagcacaga	gcttggcaca	tggtaaagagc	tcaatcagtt	acctgcttga	caatgctgac	135540
gccgatgatg	acgatgatac	ccatcctaga	ctgatgagct	ctgtaagcgg	gggtgctctg	135600
cacagagtag	acactcggtg	cagctctgtg	gaatgaatga	ggcacatccc	agaactcacc	135660
aattcataaa	aatcagatgc	agatgggatc	ttaaagatca	cctatcctaa	gtcccttggt	135720
tcacagatga	aaagacccag	gcccagagag	gtgcttggag	ctgcgcaagg	tcacacagcc	135780
aagcagctca	tttgattagt	gtcagagcca	agagctggga	gtttggaggg	aggcaagggt	135840
aagaacagga	tgctgtcagg	gaagcaggca	gggatgctgt	gttaagattc	caaatggatg	135900
cagagagctg	tgaaccggcc	agtggggagg	caagggaat	gtggtttttg	aaatggaaga	135960
ggatgacttt	agcagaggct	ctcagcccag	agggagggga	gataggaggg	ggagataggg	136020
aggggcgggg	ggagggctag	ggctgtgaaa	gtcaagagct	tattaatgca	tagagaacgg	136080
ttttaacagt	ggagagagga	aggaccggat	ttgaaagcta	cattcaagga	agtggcaacg	136140
ggatttggca	acagcttgga	tggggggagg	aggcaatgga	ccccaaggca	gaggctcaga	136200
gaagggaggg	gcaggacttt	ttgcagagaa	acaaaaggag	aggagaggag	gttagaatca	136260
agaaattctg	tgggccaaaa	cctggggctg	tgggtcaaa	gcacctgaat	tccctaggat	136320
ctctggaact	ttggtctact	cttctgacct	cccgaggctc	cccaaatgt	ggattacccc	136380
tgctcactct	cccccaaccc	cgggcccttt	atcgatcctc	tgaccataca	tctctgggtg	136440
tgtoctactc	ttgctgacac	ttcataaaaa	gaggaacccc	atttaggtgt	tttgagtggc	136500
agggattcca	agcctacccc	ctggatgggc	ctggaagaga	acaaagagca	caggccatgg	136560
tgagtccagg	tgaggccagg	gaggtgcaag	gagccagctg	gaggcctgag	ccaggatttg	136620
gggtgggtgg	agcagggggc	ggagaatggt	ggtgtcagag	gcagccgaga	aggttgaggg	136680
ggacggatct	caatgtggcc	aagaggaggg	ctcttggcac	gctcagttcc	tgtagcgaag	136740
agggcggaag	ccagatggga	gggggcgaga	acaggcagga	gcacaggaag	gtggaggctg	136800
tgggtgtagg	ctgggagtca	atgccctccc	ccaacctgag	gcctccgacc	aggctcctgg	136860
gtggcaggca	tggggaggaa	agcgtctccc	caggcagtga	gggagggaga	gccacagtca	136920
gggaacaggc	cccctgggtg	aactggcctg	agcagagtgg	atgctcctgt	tctgagaccc	136980
agacctcctg	gaacctgctg	accacagtga	tgcctgcac	aagaggggag	gacctcaagg	137040
cagtgaagtc	agggagctga	agtcctgctt	ccctctctgg	caagccctta	tctctttgag	137100
cccagtgct	ctcctctaaa	aaagtgaagt	gggtgtatgg	gtgccaaagg	attagctccc	137160
aagtcagctg	atcatcagaa	tccccctggt	agctgggtat	aatgcagagt	ccaggatccc	137220
ccactggccg	tgggccacac	acacccgccg	cccccgctg	ttaattctga	accatagtcc	137280
caaggtcctt	tctgcaactaa	tgtggcctga	ttaggtgact	ccctagcacc	aggcaggtgg	137340
gacagcgct	ctaaggggag	tagtaatgca	atgtggcttc	cttccctccc	tcccctgccg	137400
cctctggggg	tggagctgat	gccccctcac	ccaataccca	gcctagttagc	agtactttgg	137460
ttccccagg	gagctcctct	tttaaagaaa	agggcagga	cccaattggt	actgagcccc	137520
tattgtcata	gtagccacca	tttattgatg	gttgactatg	cacctgccag	atactgtacc	137580
cttaacagca	tttatcatcc	aacctcctt	tagcctgctg	aggggggttat	acataataag	137640
gaatattgta	catactgagg	aacctgagac	tccatgaggt	taaaaacttg	ctaaaaataac	137700
acagctaggg	aaaaggcaag	ctggattttg	aactaggggt	ctaagtgtctg	agcctgtggg	137760
cttcataatt	ggaccaaate	cctgtgtgct	gggcacgtgt	ccagcacttc	cctcatatga	137820
tctttatgtg	aaccatcctc	tggaaatcct	agaaacaaac	caggaaagtag	gtatactcat	137880
ccccatttta	cagatgagga	aacaggcaca	gagagatgac	tggcttggcc	aagttaagaa	137940
taatggctaa	caaacaaaaa	caaaaacaaa	aattaaaaaa	aaaaaaagaa	taatggctaa	138000
ctcatggaac	tcatagaact	ccacaaggaa	aggtgttcta	agcaccttca	tacatgctgc	138060
ttcatttaat	ctctacatta	tacagatgag	gaaactgagt	cacagatata	ctgagtgaat	138120
tgcccacggg	ggcatcagtt	aatgacagat	ccaagatttg	aaatcagaaa	ggctggctcc	138180
ccagtctcca	tacttcacca	aaccagaagt	tctgaaactc	aaactgtggg	cctgccaatg	138240
gccacactgg	cttccctggg	gaacctgtag	acatggggat	tcccaggctc	caccccaaac	138300
ctcctgaatt	agaaaactctg	ccccccgccc	caccccgctc	agagatccgc	aggggatcct	138360
aatacacccc	aaagttagg	aaccactgac	ctcaccaata	ccactttttc	cacagcaaat	138420
aggttagagg	aggcagaatc	caaateccagg	atgctatgaa	tcaaaagggtc	aaccttttct	138480
cttctgccac	ggtgcacccc	cttccctccc	ccggccaagg	ccccagcggtg	gtctgcaccc	138540
tgcttcaggc	ccatttctct	cttctgtgcc	ccactccacc	ccacccagga	tgacttggtc	138600
gcggacctgg	ccaacctgag	ccacctcttc	ctccacggga	accgcctgcg	gctgctcaca	138660
gagcacgtgt	ttcgcgccct	gggcagcctg	gaccggctgc	tgctgcacgg	gaacgggctg	138720
cagggcgctg	accgcgcggc	cttccgcggc	ctcagccggc	tcaccatcct	ctacctgttc	138780
aacaacagcc	tgccctcgct	gcccggcgag	gcgctcgccg	acctgccctc	gctcgagttc	138840
ctgcggtcca	acgctaaccc	ctgggctgtc	gactgccggc	cgcgccggct	ctgggcctgg	138900
ttccagcgcg	cgcgcgtgtc	cagctccgac	gtgacctgcg	ccaccccccc	ggagcgccag	138960
ggccgagacc	tgcgcgcgct	ccgcgaggcc	gacttccagg	cgtgtccggc	cgcggcaccc	139020
acgcgggccg	gcagccgcgc	ccgcggcaac	agctcctcca	accacctgta	cgggggtggc	139080
gaggccgggg	cgcgccccagc	cgatccctcc	accctctacc	gagatctgcc	tgcccgaagac	139140

tcgcgggggc	gccagggcgg	ggacgcgcct	actgaggacg	actactgggg	gggctacggg	139200
ggtgaggacc	agcgagggga	gcagatgtgc	cccgccgctg	cctgccaggc	gcccccgac	139260
tcccagggcc	ctgcgctctc	ggccgggctc	cccagccctc	tgctttgcct	cctgctcctg	139320
gtgccccacc	acctctgact	gcggtgctga	gatcgaagag	gccagtgtcc	gatccccgct	139380
tcccgccac	ccggggctgc	ggctccggcc	ccagctgcctc	caccttcctt	ggccttgctg	139440
cctccctttc	ccctcccagc	tcctctcctc	ccgggggagc	aggccgcctc	tccttgcttg	139500
ccccctgggc	tgtcctgact	tgtggcagcc	ccaagagggc	gtgtgtgggtg	gctcagccct	139560
gccctcccca	gttctggcca	ttaactcttc	cccattccaa	ggctgggggtg	gggcccccca	139620
ggcagccgct	gacccgcact	cctaagggcc	cacagcggac	accagagggg	ctittgtctg	139680
cagagcgtct	tccaccagca	gagcctttgg	aagctccccc	aggagagccc	accagagacc	139740
ctttggggga	tgccctcagtc	agggccaggc	tgaccctgac	ccctgcttac	cctagtcccc	139800
tcaacctcct	gacactggag	gaatactttt	ctcctaagtc	taccctggac	acttttttagg	139860
gcacctggag	agaactttcc	tctccactgt	ggccccctgcg	tggtgaagat	caaaaagaagt	139920
tgtttgggaa	aaaaaattta	ttaaaaaaatt	ctattattttt	atctactgta	agattttgttg	139980
acttgggacc	ccgaaagcgg	gatgaggtct	cagaatgtaa	ggattgcagg	gccaggaggg	140040
ttggagaagg	ggagccgtcc	cccgccatca	aagagcttcc	tggtggctgg	agggtggctg	140100
cgctcccccg	ccatgaggag	gagctgaagc	cctgcattct	agggtaggcg	cagtgtggca	140160
gccaagagt	ggtgctgggtg	gcacctcttc	tcttcatttg	tccaggggaa	gagctgcagc	140220
caacctgag	tggtctggcg	cctgaggaac	taagcctggg	gaagacctgc	tgtctggtta	140280
acagccctct	tccagaccct	gttccttccg	gaaacaagag	cagttctcct	gcaaggagga	140340
gtcacataca	cactcctggt	cacagacagc	ccaacatgg	ctttgggtaa	atgtgaacaa	140400
ggcactgctc	cctcagggaa	acacagcccc	atgccagagc	aaacacctta	gcaaacagag	140460
accaaggctg	ggtttccgcg	tacacttgcc	tccttggtta	agtgccttg	tgagtgacac	140520
agcgtacaca	cctgcacaca	gcaaccctgt	gggtatgtgg	tctctctctc	agctcctgtg	140580
aggtagaagc	catcagggat	gaaccaggtc	agagaagcag	gtttccaaac	aggctagaag	140640
agggaccgag	gaactcgggt	gatcagaggg	acaggaatcc	caaattggga	tgcatattctg	140700
gcttgaggta	caatcagaac	cttcatcttt	ctggtgtgtg	gaagagaggc	tggggactgg	140760
gaagagctca	ggctaagaag	gacttgggtt	gggatttagg	ggtgagtctc	atcagactga	140820
gcacttgag	agaagtgttg	tagtttgaat	ttggagctaa	gaatctagct	tgggcagggt	140880
gtggtcgctt	gcacctgtaa	tcccagctaa	ttggggaggct	gacgtgggag	gatcacttga	140940
ggccaagaat	ttgagactag	cctggacaac	atatcgagac	tgagtctctt	aaaaatgttt	141000
ttttaagaat	ctagtgttga	gtgggggtgtg	atgtctcaac	gtctgtaatc	ccagcactct	141060
gggaggctga	ggtggacaga	tcacttgagg	tcaggagtctc	aagaccagcc	tggccaacat	141120
ggcagaaacc	ccgtctctac	taaaaattca	aaaaaattag	ccaggcgtga	cggcgggtgc	141180
ctatagtccc	aggtaactcag	gaggctgagg	cacaagaatc	actccagcct	gggtgacaga	141240
gactctgtct	aaaaaaaaaa	aaaatctagc	ttggggagggtg	ggaatagaaa	gatagagggg	141300
gcctagatgc	tagggcttga	ggaagcaggc	tgaggttctg	tgattctggc	tagggaggctc	141360
aaatgatctt	gagaagaaga	gaagaaagga	gaagaaatca	gcattctaagc	ctgaggcagg	141420
tagactccgg	ttaagggtgt	gggtggggct	gggggagagt	gagagcagct	ggtcagaaac	141480
ccaggagct	cggagtctgg	ggtcttgagc	gggcttgtgt	caggctggct	gtgaggagg	141540
taatgggttg	gattggaggg	acagccagac	aagagctctg	gtggaggagg	ggctgctggg	141600
gcctggggag	ggggagggga	gctgctggta	aattagagcc	aggctgtcca	ggtcatagaa	141660
ttatcattgt	gaaatattca	tggggccatcg	gtccagatgc	tatttcagaa	cagtgaagac	141720
aagaggagt	tgtgagcctc	aggaagaagc	ctgaagcaaa	gccaactctcc	accaaccccc	141780
acccctccca	ccaccagccc	agacagaccc	acggacgccc	atcacgtgca	caccacact	141840
cccgagctct	cacacacact	cgcaccaagc	agagccatgt	agcacgtgca	agcacaccaa	141900
ccacccacgg	gtccacaaa	caggcagggtg	tcccctaagt	tctgacatgc	acactgacat	141960
gcacacccac	tcaatcagga	cccagcagag	atcacctcca	gcgatctcac	atgcgcagac	142020
ccccaaactc	tccaaacaac	ccagattcac	caccttgacc	cacacaccct	gagataggag	142080
ggatgttcaa	ggccatccag	cccaaccccc	accaatgctc	tgatggggaa	actgaggcca	142140
tagaaaggaa	gggatttgtc	tgagattcct	ctatccccctg	aaaaaagcaa	aattcattca	142200
cctccacat	tctgagtgt	ccccattct	gcattttcgt	ctgccagaca	ccagccttag	142260
ttgtaatata	ctcctccctt	tctctaattt	cctgcatcta	ttcagttacc	cagtcctcca	142320
cccagccaca	gtctatccct	tccttcccat	tctccccacc	acctccctgc	tccagctact	142380
cattacctca	tgcctggaat	ataaaagaaa	actgcgataa	cctcctcgct	ggtttcctac	142440
atggaatctc	tcctccctc	ccaccagcc	ataccgtgggt	gaccagattc	atctgatcaa	142500
aatttgcata	tgttatgat	tcactcagga	gcctgtaatg	gcttcctaag	gcttaaggg	142560
ttaaaggtaaa	acaccttagc	agagcatcaa	agatccctca	gagtcgtgta	ccaactgctt	142620
ttctagcctt	ttctctcaca	atctcatccc	aaaccttcac	tccagctaga	acgtttgtat	142680
catactggcc	accagttatc	atgtatgtga	aaaccaacca	ccgactttga	gtgccccctt	142740
aaaattttctc	agtctctcct	gaagtaggaa	acctcttccc	cctcctcaga	tctcagactc	142800
cagagccctt	tcccaaggcc	aagactgcac	ctctctgacc	atatacaggg	gttcttcaaa	142860
gcagcagaca	gaggctcagg	ctctggctcc	ctccaagcag	acggctgccc	ccgactggcc	142920
accttgggaa	gcacagccag	gtttcagtcg	tctagaacag	agaatgagca	tctaaccgcc	142980


```

tggggagagg actaggacac cagatgataa gggtttataag cccttaagcc tctaaggttc 143040
ttacacccag agtagggggg ggacgggttct cagccctgtt tccctagctg cgggctccca 143100
attttcgatc cctaataccga gaggaactcc tctccaatga aatacagact tgggactctc 143160
aggacactgt ggaagggaaa tttcccaaca gactctgaga gtccaggagg ccagggatag 143220
accaggtggc aggcccaagg tccagctggg gtcagggttc tatatgaatt tttaatgctt 143280
ccagatagac ttgtcagatg ttctgaaaac tgagcatctc ctttcacctc tgtacatgat 143340
gcccttctcc aacccattg cccctgcagg agggcaggcc tgggacagat attcagtggc 143400
ctctggagaa acggttttgg gacagtagaa gggtaaataa cctagttatg ttcccactag 143460
taagctgtgt gaccttgggc aagttactta acctctctga acattagagt tctgtgggtt 143520
tgtttttgtt ttgtaagctg gggacaatag tgccagccta aatcaatttg ttgtggggac 143580
tcagtgaat agcccatggc aaagtgcact acatgcttgc tggtattatt ctctttcctc 143640
aagttctgcc tccctcttcc agcttttctt ccaaccccaa agatgtctct ggctattgct 143700
tcgaaggtag gaactttggt tggttctccc ctttctcttc aggcccaaac tccccacctc 143760
aagatccttt ggcctttgta gaaacttcag gtgaggaggt ggcagagaaa taagaaagtg 143820
tgcaaggctg gtggagtga agaggaggat agatggcgaa gccctagcag aggggaggga 143880
agtgggcagt ggagagagg 143899

```

<210> 16

<211> 215980

<212> DNA

<213> Mus sp.

<220>

<221> modified_base

<222> (1001)..(1100)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (2123)..(2222)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (3728)..(3827)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (5168)..(5267)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (7481)..(7580)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (8849)..(8948)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (10375)..(10474)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (12270)..(12369)

<223> a, t, c, g, other or unknown

<220>

<221> modified_base
<222> (13438)..(13537)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15902)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (15939)..(16038)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (18223)..(18322)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (20974)..(21073)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (24403)..(24502)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (27574)..(27673)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30892)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (30901)..(31000)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (34443)..(34542)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (38205)..(38304)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42373)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42386)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42393)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (42461)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (44809)..(44908)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (51380)..(51479)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56740)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (56765)..(56864)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (62818)..(62917)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68518)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (68534)..(68633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (74552)..(74651)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (81446)..(81545)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (88519)..(88618)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (93791)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (93794)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96565)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96570)..(96573)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96579)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96590)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96596)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96602)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96616)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96629)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96633)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96668)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (96715)..(96814)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (104447)..(104546)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114521)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (114527)..(114626)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (127063)..(127162)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (139133)..(139232)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (151051)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (153242)..(153341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164706)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164708)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (164710)..(164809)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (182242)..(182341)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192158)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (192192)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (198842)..(198941)
<223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (199437)..(199438)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (208276)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215974)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215976)..(215977)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (215979)
 <223> a, t, c, g, other or unknown

```

<400> 16
ttggggggtat aaaccagaa gtgggattac tgcaccatac aataatcctc taacttcaag      60
caatttttcc acaatgggtg tatcatttta cattcccact ggctacgaga aggggtccca      120
cttctacaca tcttcaccac catttctgtt tttgtttttg agtaacagct gcctaatagac      180
tgtgaagtgg tatcttatct cagtgttgat ttgcatttct ctgatcatta atgtgggaag      240
gcatcgtttc atatgtttat tggctgtttg tgtatcatct tctttggcga tgttgattca      300
agttatttgc ttgttttttt aattggagtt ttaaaaaatt gttgttgagt tgtgggagtt      360
cttcattagc tctgcatatt aataccctga tgaaaatgat taacaagtat ttgcttccat      420
tttgggggct tccattctgg gctgttttta ttcttttgat actcttttga ttctcaacag      480
tttaattctga ctaaaattca gtttatttct tcttttaatg gccatgctat tgacacatcc      540
cgtaatcact gccaaatcca gtcatgaaga gtttctttca agagatttat agtttttagct      600
ctttaagttt gtcattgtctg tttcacttaa ttttgtatag tgtacaaaag tctaacttca      660
ttcttttcta tatggcttgc tactagtata cgaagagcta aatttctctt tccttgagtc      720
tcaacctctg atgtgtagca atttcttcag aggaaaacat ggtgggaagt tccttaaaca      780
taggatgctc catggagggtg aaatagttca tcctacaggg aagcttggtt aacacaggaa      840
gtacatactc agcagctcta gtaagtgaat gaaactgact ggaggcacta ggtccctcct      900
tccctacgca tatagaagct gtaaggattg ggaagagata ctgtcaggtc agctcagctg      960
ctgccccgaa gaagctcaga cccactggcc tggctccaag nnnnnnnnnn nnnnnnnnnn      1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      1080
nnnnnnnnnn nnnnnnnnnn atcactcttt actcaggcca cctacacgct gtttatagcc      1140
tgctttgttc tctttggcta tacttctgtt ttatgtctat gcctcccttc tttcttttct      1200
tttctcttct cttctcatct catctcatct ttcttcaggg gggagcctgg tctagaactc      1260
acaaagattt gactgtctct gtctccttgc actaattaaa aaatctttta caagcatctt      1320
ttagcaattc ttacagggaat attttggaat gttaaactct gattgttagc gggctgaaga      1380
taacaatagc tctgatgata aattgcttgc caggcaagtg tgaaaatctg agtttgatcc      1440
aaaaagccgg gtacagaggc caaagagtcc ataatcctag taggggcagg aatcagggat      1500
gggtgggtcc ctgggggtttc ctggtttgtc agcgtagccc aattgggaat agccagggtt      1560
cagtgaacga tgctttctgc aagctgagag aggtccttgt tcaatctctg tgacccaact      1620
ggagggagaa gagagccagc tctccagaag tggctcctctc aactttgtgc atgcatgtcc      1680
atgttcacac agggaatgga taatgcttaa aaggaagacc ggcagggggg tggtaatgca      1740
cctccttttg tgacatgctt tctcttgtt catgctgtct cagggtgtgg cggcagcacc      1800
aaaaaccagg tgtatgtttg taatcccagt attctctggt cgtcagtagg aaatgaaaag      1860
cgaggtcatc ttcgtataga gttagcaaac tctaagccag cctcggctac atgagacttt      1920
gtctcaaaaa aaaggaaaaa tcaaggagga cggctcccga gcaactgtcac ctgaagctga      1980
cctetggcct ccacatgcat gtgcgcaaac acatgtcctg cacaaacaca cagacaccgc      2040
catctgctcc ccgacaaaag aacctgaaac cagtatactt tgagaatttc ccattcatag      2100
ttaccattgt gtgttccttg tgnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      2160
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      2220
nnggtggtgc ctttctctta cccagtctag aagggtctgga ggcaggggtg atggggcact      2280

```

ttgaactccc	acctaggcaa	aaaaccagtg	gatctctggg	ccagtggtgt	gtttgcaagg	2340
gaataaggta	gagagccgcg	gaggaagaca	ttgggggttc	tatgagtatg	tgaaggggtg	2400
cacacaccac	acacacacat	tttttttggt	ttaaattttac	aaacattaaa	ataggctgta	2460
atgtggctca	gtgggtagaa	aaacctgctg	tctaagcctg	gtacgagttc	aatccctgac	2520
aagctggaag	gagacaacca	accacaactg	ctagcagcca	gaagcactgc	ttgctaacac	2580
tcaagagagc	ctggagtgga	agacactgga	tccccagcag	gcaagcctgc	aagaagatgt	2640
gccttgcccta	gacaacggca	gaacaaacat	caaggctggc	agagctgtcc	aggactgttc	2700
atattaatca	tgtatagata	agaggggaatg	gcacagacag	aacaattcaa	cacacgggggt	2760
atgaaaggaa	aggaacaagg	cacacaaagg	acaaayaacc	tagcatacaa	gaaagcctaa	2820
gcagagagtg	gcacttccca	gaaggggagtc	ataaaataga	ctgaattcat	taaaacaaga	2880
gccaaagata	aacggctcaa	aaaactcacg	gaaaacaggt	caaaataacg	tcacccatct	2940
gacagttgat	actgtcaact	taaccgtact	tagaactcca	gcaggcacat	ctccaggcat	3000
gcccctgaag	gggtcttttg	actaggttaa	ctgacgtggg	agtgcacca	tctatggacc	3060
gaagcctcag	acagaataaa	aaggagccag	tgagctgagc	gtcagtgctc	attgtctctg	3120
gcttcctgtc	tgtggctgca	gcgagacacg	gtgcttctctg	cttttagctgc	catgacagac	3180
cacaccctca	aaccgtgaac	caaaataaacc	tcctctctac	attgctttta	ccaggcattt	3240
ggtcacacca	atgagaaagg	ttaactaata	cagcactcaa	tacttaaaaa	cataaacacc	3300
aaccttgttt	gcattgttga	gactttgaag	ctcacggggc	agttatgccc	aatgccaggt	3360
ctgctggcta	aggggtgagag	tgcacaccta	taatccagc	tgctgtggaa	tcagcaaaag	3420
cgctacagat	ggaaggcagc	cagggcagct	gagactgact	caaactgata	gaggtgggag	3480
gcatagagaa	aaccagatta	atagagtgtt	ccccactatg	caagaagccc	tggtgttcag	3540
gacgagagaa	ctaagaatac	agaagtctac	tgtgtagaag	cactgctagg	tcacacagaa	3600
acatcactca	agtgtctctg	gatgtacac	ggagggcgtg	tgaagtattg	cttcctgatg	3660
atctgtatct	actacagcac	tgctgtttta	gtatgcgctc	ctccactaca	gctcctcacc	3720
acaccaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3840
aaacacacac	caccagttag	agaaaagttaa	tcaggccgaa	tggtggcttt	cccctgtatc	3900
caggctaccg	tcaggacggc	tcactgccac	tggaactct	gcctgaacaa	agcccgcagc	3960
caacgtgggc	ttcaggggct	ctaaacactg	caatcaaagg	ttgtgtgtgg	gggtgggggt	4020
gctgctgcta	ttcaaggatt	cccaaagctt	agatgtattc	atcatactca	caggaaagcg	4080
tggtcaaccc	atcactcatg	agcagtcggt	accgggggtga	cctattccct	gtagaaatgg	4140
gacggatggt	ctggaaaaagt	tgacagaaaa	gttgattcat	taggcaggct	ctttgcccaa	4200
gccctgaggg	taagcaaaag	taactggcag	gagactaggt	ttgccattaa	tctgagacaa	4260
gatgaaccac	ttgcccatcc	tcctgacacc	taaatactaa	tgaaagaaca	atggattgag	4320
ctggcattat	taaaaacgat	agaaacagaa	gtatcaatag	tcattgtgtc	tttctcccat	4380
atgtcaaaac	aatgtgtaag	atggcatcga	acacatgcag	aaactgttta	gggaacatgc	4440
tgaaaaatatg	aagtaaaatt	aaaattggaa	agaaagacaa	tttgcctaaa	gcagctcaga	4500
gctggagaag	ggaccgaggg	agagataaca	gcaacgtgtg	gacatacgga	tctggggcag	4560
agcagtcacg	gactcagccg	gaaagggtgg	ggcagcctct	gaaggaagtt	aaggtaataa	4620
gagccacaag	gtgattggcc	caggagtggg	gccaccttca	cctcctgcct	caaagtctga	4680
aggaatgatc	ctggagtctc	ccatctattg	atatatgaaa	ttcacagtat	gttttagaac	4740
ccactgaatg	atgggtagat	taactaaaag	aaatttaagc	gggggtgggtg	aggtctttta	4800
atcccagcac	ttgggaggga	gaggcaggtg	gatctctgtg	agttcgaggc	cagcctgggt	4860
ccaggacagc	cagagataca	tagagaaaac	ctgactcgaa	aaaaaaaaat	taaaagctca	4920
tcaaaaacaa	aacaacaaca	aaaaaaaaca	aaaaacaaaa	caaaaacaac	aaacacccta	4980
tagtacctgt	tggtgagttt	gagtgaagtga	gtgagtgtgt	gttagagaga	ggggcgggga	5040
aagtgtgttc	tggaaatggg	agaaagagaa	tgtgcatgtg	tgtttctggg	atgtagacaa	5100
aactacatgt	cttccatcaa	atgcaatgtt	taattatcta	tgagttgaac	catcttcatt	5160
ctgctaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5280
caaaccagta	aacaaaatcc	tgtaagataa	agcctaagac	aagacacttc	ctggggctgg	5340
ggagttgctt	agaccataag	gagttcataa	tccaggcgtg	agagcccag	ttcagggtccc	5400
tgggcttcca	agtccaggag	agaccaagga	atcaacaagt	ctcgactttg	gtctctagtc	5460
ccatgcacac	acacgcgtgt	aaatacgtag	atgttctact	acacacagaa	gactgcacct	5520
ggctctctca	catctcagcc	aacatataaa	gcctgcatta	tcagaacatt	ctaggttcta	5580
gtttcagtc	actcttacac	agaatggcca	tcataactccg	tctacaactt	ctcctgatct	5640
accacagtgt	cattgcttca	gtattaacaa	aaccagaat	aaccagctgc	gtagatcctc	5700
cctgatgccc	cagtcatgtt	cttactgaga	ctactaagtc	acaaggtagc	actctggatc	5760
caaaaagcaa	tatccaattg	agagttacaa	cctataagga	ggagtttacc	ttcattatag	5820
ggcactggat	tcccaatctt	taatccaacg	tcttcagcag	atttcataac	ttccaagtcc	5880
atcaaaaaca	ctactttcct	acaaagacag	acacaagtta	gaattaagaa	ctctgcagcc	5940
tttcagatga	gttactaaga	agcttacttt	agtagttgtc	tggtctaaac	tgtatccttt	6000
accaaccttt	tctcattctg	gactaacttg	agaagtatta	attcctaagt	aaatacttca	6060
cttattcttt	ccccacatct	ccaatgtttt	tgtctttaat	ttattatagg	gcaattcatt	6120

tcctatctag	ttccctgatt	aaaacagtag	accttgctgc	atgccattat	cctcatggag	6180
gcactgatac	aatttagatt	attaaataca	aaaccctaaa	acacaaaaag	atgatttttt	6240
tttaaaacaa	gatttttaaaa	aaagcatgtg	ctacgcttcc	ttctgccact	aagcctacac	6300
atggctcctc	gactgaattt	ttcccctcat	tctgcttcat	ctaataatgtg	cttttcaaac	6360
ctggaattga	accagggact	tattcatgct	aggcaaatgc	tctaccatag	agctataccc	6420
ctccaaactc	catctcaaat	atcattttcca	aagacatttt	cttgggtctct	tatttagatc	6480
aggtttcttt	gtcctcctgc	agctatgact	tcattccttc	agaacactcg	tcttagcttt	6540
aagttctgta	ttaattagt	attgttttca	ttctctctgc	tagaatgcac	tttcaataaa	6600
ggcaggtagc	cagccacagt	gcttaattaa	gcaacagccc	aacgatgtca	ttcactacat	6660
actgggacaa	gatgcctaac	atcatctgca	gataaagacg	aactactggg	gtcaggagac	6720
agctaagggg	tccagggcct	gggcacgctg	agtgtgagca	ctggagtccg	ggtgcccaga	6780
aacgcacata	aatgcaatat	ggatgtggca	atctacctct	aattccttct	ttaagacagt	6840
ggctctccag	agcaagctgg	ctagcaagac	aagccatata	agtgagctct	gggcttgacc	6900
aagaccctgc	ctccaggtgt	aactcccaag	caaaaggatg	atggctcaca	aatctcaggc	6960
tatcatgttc	atgtacaaaa	tgtcaaccgg	catacacaca	tgacacacac	tgaaaactgg	7020
gagaaaaata	gaagaattgc	aacaaaaaaa	tgtaatttga	ggacacataa	ttgcaggcgg	7080
ggagtggggg	gatgacagaa	ggtgaactga	gtggaccgag	ggaaagctgt	gctagcggca	7140
atgagaagaa	gggtggggca	gtctgagcaa	gggttcagca	atcaccacgc	tttactgtct	7200
gcacagcctg	gctgtagaat	gctgggcttt	atcacacaga	attattcagt	atgtgctatc	7260
tttacagtaa	agttattcta	tcaggctatg	ctacttcaat	agaacaagcc	tgaaaaagt	7320
gtctgctgct	gagaacctga	caaagatgac	ctgttagaag	tgtctgcaa	gtgtggaatt	7380
ccagcactgg	ggaccaggag	ctcgagggtc	acccagatg	cagggagtta	gaggccagtc	7440
ttggcaacat	aacatcatgc	ttcagaaatt	aaaaacaaaa	nnnnnnnnnn	nnnnnnnnnn	7500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7560
nnnnnnnnnn	nnnnnnnnnn	catgagatag	ttaataaact	gaagaaagcc	atacaaggag	7620
taaagtagat	agttgcaagc	atgaagaaag	acaaaccact	tgagcttttc	ttttgtcgt	7680
aggaggaaac	cagacaggtc	cagagagatg	gctcagagat	taagagcact	gactgctctt	7740
ccgaaggtcc	tgagttcaaa	tcccagtaac	cacatgggtg	ctcacaaaca	tctgtacagc	7800
tacggtgtac	tcataatacat	taaataaata	aataaataaa	taaataaata	aatcttaaaa	7860
gaaaaaaaaa	aaaaaccta	ccaatcagcc	aggcgatgg	gacacatgtc	tttaatccca	7920
gcacttggga	ggcagagaca	ggtggatttc	tgagttcgag	gccagcctgt	tcttcagagt	7980
gagttccagg	acagccaggg	tgatacagag	aaaccctgtc	tcaaaaaaca	aacaaacaaa	8040
caaacaaaca	aacaaaaaag	gaggaagcca	gacaggatgc	actttatacg	tgaatggaat	8100
tgacaaaaga	caagttctat	aagtgttagg	gaaaggggga	ggacaacggg	ggttcatgtc	8160
tgtgggtgga	cacgtattag	aaggctctgg	gtatcctgtt	tccgacaaac	aggcactccc	8220
aatcacacag	gccactggat	gtctcaggca	gagaaagatg	tgatagattg	actttttaac	8280
aatcacagac	tgtgtggaaa	atatttgtaa	ggttgtcatt	gtcaccagag	atagagctga	8340
tggttatttc	aacgaggatg	ggacaacaga	aatgggagag	agggatgtga	gaaccatttt	8400
gaaccagggt	gatttactgc	gcacgtgtat	agggcttaca	gggagtggga	tatgtagagg	8460
aggcctatgt	tcctaacttt	ggtaatgagc	ttattacagt	tactatgcac	agcctggaag	8520
atactggaaa	aggtgcaggc	taggctagaa	aggtactaac	tgaggggttg	acagcccctt	8580
ggatgtcagg	atgcagcaag	cctacctctg	tatgtagtca	atcccttctc	aggctatggg	8640
tectgcagat	catcctctc	tgtatccatt	attcccagtc	catcctctga	gtggctccct	8700
cttatccagt	ttaacaaaat	gctgactgca	agctcccaag	cccagggtct	tggtcctttt	8760
actccttggt	attgtacttt	accctgtttg	cttgggatag	agtgtgccct	ttataaacat	8820
ttgtgaaagg	gggaatgaag	aagaataann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9960

gccgtctttc	aggggtgttga	cgatgaactt	gtggacctgg	cagacacagt	tgctggccag	10020
ctgccctccc	tcgaccaggg	tgttcagctg	cgtggccagc	atgaacgctg	caaaagcaga	10080
gagagagggg	ctcagtcctc	aagcctttcc	ttaacccgaa	agctcatcac	aaggagaacc	10140
attaaataca	gctgtttaaa	actcctccgc	cctgcagaga	ggaaagcagc	atcaatccgc	10200
cccatgtaaa	agtctgaggc	tcttcctaaa	tggtatctgt	ttctcacagt	ctccaaatca	10260
tttttactgt	aattctagtt	tctggggaaa	gacctttctc	ggtcttttagc	cccgtgacta	10320
gagacaacag	gcaaatattc	cagaaaggcc	cccattttct	ttttaagct	tctannnnnn	10380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10500
ttcggtcctc	cgaatttggg	tttcttctgg	gacgtggtag	catgtgactg	tcactccagt	10560
gcttgaggca	gcagaggggt	caggaaactc	aggctggcat	tagctgcaga	gctggagcag	10620
gtcctggaga	acagaaactt	tggttgagc	attaatgaac	tagaagaatt	tttttgtctt	10680
ctgttaaata	taaatacctc	cattatcttc	tcataaacag	tggtgccttt	ttattttaagt	10740
ttttaaggat	caggcacaga	gactccatgc	cagactacca	ctcaaccact	gagctacacc	10800
cccaacttgc	ctttctgcta	tttttttaaat	tgatcagtg	gccaccaaac	atggggagag	10860
gtcagggggc	tacgtggagg	aattgtttct	ctcctaccaa	gtggggccca	ggtttcaaat	10920
tcaggtgacc	tggttgaggc	gcaagcacct	ttacccttaa	gccatctcat	tggettcatc	10980
ttttaatggc	cccttccctc	gctctgaggc	aggctctccc	tatatagccc	tggtctggcct	11040
caggctcgca	ggtccaccag	tgagcaccag	gtttctgctt	gtccttacct	ccccagcact	11100
gtgggtataa	gcatgtgcca	ctgtgtcaaa	ctcagtcact	aagctttgcc	aagccatagc	11160
ccagcccttg	agtttactgt	ttgtctgtgt	ggtgatttgt	caaaccactt	ttgttccact	11220
gaggtatttt	gtcaagtttg	acaaaattag	ctagatcttg	aggctctttt	ttctggaatc	11280
ttctgttata	gcttagtctg	gtcttgaact	catgatcttg	cctcaacctc	acgattattg	11340
aggattattg	agatggacag	gctgtgtgac	catgctcggc	tgtgtgtttt	agcatgcatt	11400
agtcatttga	aaaacgtttg	ctcatgacac	tttacaggtc	ttccatgttt	gatatgtttt	11460
atttaaatcca	aagtaattcc	agcaccagag	gctgagacag	gaggatctca	aggtcaacct	11520
agagatgcat	agcaggcggg	gccccactcg	gttaggttaa	tatcatcact	gacttcagga	11580
gaaaagtctt	aagtattggg	gactaaaagc	aggaggatct	gaagttcaag	gtcatcttta	11640
ggaacttagc	agacttgagg	ccagcttggg	cgctgtggga	ccctgttttt	aaaccagaaa	11700
acaaattgaa	aggaaaaaaa	aaaaaagctg	gaggaagtga	atgtgagtgt	tcacatagtc	11760
ctgtttccac	aagaaaacag	ggttactttt	ggcaacaaat	aggtgctttc	tttgaaggct	11820
ggcatttttg	tgacttgtca	ttggagaaat	gatttaatta	agacttttct	actgagtgcc	11880
tctgaagagg	ctctttttaa	tttagtttaa	ttttatctca	ttgttagtgt	ggtgtgcttg	11940
tgcacacaga	aggcagcttt	ctagagctct	ttcactctct	cctccacagc	tcctggagtc	12000
aaactcaggc	cctggctagg	caagctctta	ggacagtgtt	agctgtagct	tattaagttt	12060
ttaagaattt	ttataagact	ctgtttttct	ttctcaggtc	atgatacagc	aggaaaatac	12120
atccataaag	cccatcctgc	aggtcattgt	aagtaccggc	atgtgtgttt	agcataatga	12180
agatgggtca	cttatagtta	attaaacatt	ggattggatg	gaagacatgt	agttttgggt	12240
acttcccaga	aacacaaatg	cacattcttn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12420
cttctattcc	ttcattacat	tactgttaat	agtggtaact	atgtaccaaa	gagtcaaatc	12480
actcttgga	catccaaggc	agaaggaagg	ctggcaaaaa	tgtagatga	ctgggatgg	12540
gaatgtactt	cagtttgtag	aggaggcctt	tggttcattc	catttctggc	aatgcataga	12600
cctgtaggat	ctcagcactg	gtgggggggt	gggggtgagg	gtgaaggggc	gggaggttaa	12660
aggcagaata	gtcataaatt	caaagtctgg	gtcctggaaa	gaggactaaa	cgattaagag	12720
cttttagctgt	tcttctagag	aacctgggtg	gatccccagc	acatggtgcc	tcacgactgt	12780
ccgaaactct	gattctaggg	ggatctgaaa	accctcttct	gccctctgta	gatacagaac	12840
acacatgggt	cacatacata	catgcaacct	aaacaaccca	tatacataaa	atattttttt	12900
ttcaaaaaga	cattcaaat	cttctctggc	tatatagtgt	ttaccaaac	tcaaaaacaa	12960
aacaaaacaa	aacaaaacaa	agaatcatta	atgttttgcc	ttcatgtatg	tctgccacc	13020
acggacatgc	ctggtaccca	gggagattaa	aagaagacat	tagctcccc	ggaatggaga	13080
taggtatgat	ctaccacttg	ggtgctggga	acctgggtcc	cctgcaaaag	cagtaaatct	13140
ttttaacccc	taagctgtct	ctcccaacgc	ctaaagattc	ttgtaacaca	gcatgatgag	13200
cactggcaag	catagcatgg	taatctgact	tcagggcggc	agatttttag	cttaatgctt	13260
gattattaga	agtaacgtac	tagatttaat	gcctggagct	tcaagcaaca	aaattaactg	13320
aagaataaaa	ataaaaaccc	tgccagccat	gatggtaatc	ccagaacttg	agaggcagag	13380
gcagggtgatc	tctgtgtttt	gcaaggccag	ccacaatcta	catagcacgt	tcagtanann	13440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13560
agaaattcaa	agaaacaaat	gccaaataaa	tacacatatt	gtaataaaga	gataattgtc	13620
taaaaaactc	aaggctttaa	atggtaagat	atcatattct	tggtatgaaa	gatctaattgt	13680
caaaatatat	caattttaatg	caattatgta	tattcaggag	atctctgggt	ggcttttgaa	13740
cttgatagca	ctcttataat	tcacatagaa	gaaaaaaaac	catgaaaact	gccaaacatt	13800

attagaatac	tccacagatg	gtatttttggc	agcacatata	tcgaagggct	gtgaaagatg	13860
tgtagatcat	ccacgccttg	ctagggagag	ggcgggtgtg	tgtggggggg	atagctgttt	13920
gggaaaataa	cctggtaatt	cctcattagt	taaatcatag	tcagaacctg	gactagcaac	13980
ttctctctaa	aatacattca	ccctcagcat	ctgcattgcc	aggaaaccac	tcctagcagg	14040
atctgtacgt	ggatcaaggt	agtagcatct	gcatttaatt	gacattctcc	taaatgcttt	14100
aaattatctc	tagattactt	atagtagcca	agatgatgca	aattatgtta	cactgtatta	14160
tctggggcgt	aacaagaaaa	tgtctctact	cagggttcatt	cagggtgcagt	acttcccctg	14220
aatactttctg	aatacacgga	tcaagaagcc	acagaaagag	ggctaaccat	atacaagcat	14280
atagtagact	aataaccatg	tacaaccata	tagtacacta	atattcagtg	cattactcaa	14340
aatgcaaaca	gatggaaaaca	atccaacagc	ctgtaagctg	aaaaacaaga	taagcaaaat	14400
gtgctggggc	tagaggccca	ggctctataat	tccaactaag	gtcgaggcag	gaggcatctca	14460
agttcaaggc	cagcctagac	aacttagcaa	gaccttgtct	caaaaacaaa	agtaaagagg	14520
ctgaggatat	agctcagtat	agagcatctg	cttagcatgt	gcactgacag	ccgtatcaca	14580
gaggaaaaaa	aaaaataagc	aaaatgtgat	ctgtctgcac	aacaggatat	cacagccccc	14640
taccacaggg	gaacgacaca	gtaacacaac	aaaaacttag	ccctgaaaat	actatggtaa	14700
ataaagaagt	gtcactgagg	atcaggaaat	gcactgactc	atttacatta	tatagaaatg	14760
agaagatcag	tgagcctcta	ggactcaaga	gatttgggat	tggcagctaa	agggtagctgg	14820
gtttctttat	gggggtaaga	aaacattcta	aacttaactg	tgagaatgac	tactcaacaa	14880
tgtcaagtgt	tcaaaaatca	tacttttttt	tttttttggg	ttttcaagac	aggggtttctc	14940
tgtgcagtc	tggaactcac	tctgtagacc	aggetggcct	cgaattcaga	gattcacctg	15000
cctctgcctc	ccaagtgtctg	ggattacagg	catgcgccac	cattgtccgg	ctcaaaatca	15060
tacttttaaa	aattgccacg	tgactcatga	atacaatcag	aggcgggaga	gtcagatggc	15120
aaactcagga	taccagtgtc	ttttatgtct	gctgcccaac	tatcaatttc	ccatagttac	15180
cagagaactt	tttggtttgt	ttcatcttat	ttgttgcttt	tggtagaatc	tcaatatagt	15240
aagatacaag	gctggcctca	tactatatag	ctgaggacga	ctttgaactt	ctaactctcc	15300
tgcttccatc	tcccaagtgg	tgggattaca	gggggtgtacc	gctatgccc	gcaagcacia	15360
agccatttga	accacacccc	agccttttca	gagaaacctg	tacaagcctt	agtgcccttag	15420
catattaagg	caacaaaaga	cataatgcgt	ggctaccata	gagtgtttgc	ctaccatgtg	15480
tgaggctcta	ggctaaatgt	ccagcactta	taaaaaagag	ttaaaaacac	tcatgactca	15540
aggatgacta	tgacgtcttg	tgtacaaagc	cccgcatcca	atccccagca	ccgtgcacat	15600
caggcaggct	ctgtagagga	cccagcttaa	ggctcatcct	aggtaagtta	gaggccttag	15660
atggctacat	tagatgagac	cctttctcat	aaacagaata	aataatttaa	agctcctgat	15720
caaaactcat	gccttcccac	caacactcaga	ataaagcact	ctactggccc	tttaaggact	15780
gcccactctg	aagagaaaacc	taagttacat	tccttgcttg	tgtcatatgt	gataacaaac	15840
tactggaaa	tacgaaaata	cagtcttaag	cttggtcaga	aagcttcccc	agcaacatga	15900
tntcagagga	cataatgcag	aaagtggaca	aatgcaaann	nnnnnnnnnn	nnnnnnnnnn	15960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16080
atagctttca	gggatcaaac	tcaagtgtgt	actgagcgct	tatgcccagt	gcgccatcgc	16140
accaggcctg	cttctttgtt	ttttatgggt	ctgaatcaat	tagcaccatt	acaacaatgt	16200
tgacaatcag	caagtacctt	tctctacctg	gctagtaaga	gaagtaagt	cctttgggtg	16260
gtgaacgcag	tttctcttgt	gaagtgcagt	gacttgatct	ttgctcaca	cgtttttttag	16320
gtccttaagt	tgcttgggtt	ttatgggaaa	ggctcttggg	ttttttgaaa	agattttact	16380
acaacttgat	ataatcatta	tttttaatcc	tttaaatagt	atgacttatt	ttaacagatt	16440
aatattgaac	tgctctttca	ttcttacata	ataaatcctg	ccttaaaaaa	taatcctctt	16500
agcttccttt	ctctattttc	aaatttgttt	tatatttttg	catgattttg	aacattttata	16560
aaagtaggca	gacaacacag	tagaaccaag	tccccatata	gctgtgcaca	tagcttcaga	16620
ttattgctcg	ataataggtc	ctgtttttgt	tctgttttct	caaggggatg	tgttattgtg	16680
tgtgtgtaca	catacatata	tatatgtatg	tatgtatgta	tgtatgtatg	taatgaactc	16740
cctttaacaa	aacaagtact	gggctgggaa	gacagcacag	ttagttatgt	gtttaaccgc	16800
acaagcatga	caaccagagt	tgagatcccc	accaaccgca	taaaaagctg	ggcatagtgg	16860
cattgacctg	tagccctggg	gctggatgaa	agctggggag	gcaggtagat	cggcagagct	16920
tactggcaac	aatctgccc	agtaggtaag	ctctgggctc	agacatccta	tataggaaaa	16980
agatgaagg	cgaggcgag	cggcacacac	ctttcgtggg	agtgtctgag	aggcaggggc	17040
aggccagtct	ctgtgaccag	cagcctggcc	tacatgtcaa	gttgaggag	agccagagcc	17100
accacctact	gagactgtct	cagaaataag	ttttttaaaa	aattgagatg	aaggagctgg	17160
taagatggct	tagaaggtaa	aggcacttat	cactaagcct	gaagccccga	gtttgacctt	17220
ggaccccaca	ctgtagaacc	aactcctcca	agttctcttc	agacctccag	agagcacaa	17280
gtgtatgcag	acacacacac	taagtaagt	aattgtaaaa	acatgacgta	gtggcactgg	17340
cctttaaaacc	cagcattggg	aggaagaagc	gggtggatct	cttgagtttg	agaccagctt	17400
agcctacata	aggaatttca	aggcagccag	ggctacctag	aaagtagctg	tttatgaatg	17460
aatgaataga	aggaaggaag	aaagagagac	agacttaaaa	aatatatgct	ggagagtaac	17520
agaagaggac	accggcttgc	tggtgtcttg	acctctggct	tgtacacata	cacatgtgta	17580
gtgcatacac	ccacatacaa	ttgtactcag	acacacaqaa	acatgtactc	attcatatac	17640

tgcacacctc	aacactcaga	aatgaaaaa	acaggtacca	tttacacctc	cgtgttcggt	17700
ttccaaccac	tcatatgtat	gggttgtaaa	tgcttatatc	tgtatgtgtc	tgtatatttg	17760
tgtatacatt	caaagttgag	tcaggatcca	acgtaaactt	ggatagtagt	gggttgatgg	17820
tctggaagcc	tgctcgcagc	tgtctttttc	tcctcgtacc	ttttccctcg	tttgtttcta	17880
cgacagcagg	tcatttgtct	ctaagtgtta	gtttcccatc	ctctctcttt	tgctgatggt	17940
agccttgtag	tagtcacctg	tgttctctgt	aaaatggcct	tgccgtgtta	tttcaatatg	18000
ctatcatcct	catcttgcta	tattttcattc	aatatatgta	tatattacaa	gatagattaa	18060
aattattttta	attttatgct	tatgaatggt	ttgcctaagt	atattgcacc	ttgtgtgtct	18120
agtgtccaca	gaactcagaa	gaaagtgtca	catattctgg	aactggaatt	gcaggtgggt	18180
gtaagccacc	atgtgggacc	tggaaccaa	atccaggcgc	cannnnnnnn	nnnnnnnnnn	18240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18360
cagacttcct	gtttctctca	tggaacagtt	tccccctgag	atttactagt	ggaagaaagg	18420
cactcaaaaa	gcagggagcc	ctcgtacaaa	tgagacttcc	tagctatata	attaggccga	18480
gatgcacaca	tccacagtca	ttacccttct	tcagagcctt	tgctcatgtc	agtgtatttc	18540
gccccatgtg	acttttagaac	tggtctgttg	tgttccataa	aaagtgtagt	tggtgctctg	18600
attgggattg	tgttaaattt	atagatttct	ttaggagag	ctgacaactt	tacagtatta	18660
aatgttttca	tccgcggaaa	agggtgtctt	gccacttact	tggtgcttct	tttatgccct	18720
taagtaaagg	tttatagttt	tcttttatatc	agtcttgcac	atttcctggt	agattttattt	18780
ttgcttgtaa	ggtttccttg	gttggtattg	tgtataaaca	ttttctccca	attacatacc	18840
ataattgatg	gttaggagta	taaataaaaag	gtagaatttt	aaaattgctg	tatgaatgac	18900
tgctttgcta	taattcaaca	tcttttcatt	tttatgccaa	tctacttgac	atgttttaggt	18960
taaatgatga	tatctgtaca	gagtaatcct	ctgatccagt	atttgcacat	cttactttct	19020
aacgtccata	gcatagatac	acatcttata	ctgttgagta	catatatatt	taaggatattt	19080
accatagtct	tataatatgc	agcgtgcttt	ggttcaagac	agttgccctg	tgttcctcaa	19140
cattaacatt	tttttcatca	caaatacaca	ttgaccttta	tcaaattttt	aaaactatct	19200
tgagagaaat	gaccattttt	cttaatctgt	taagtataaaa	tttttataaa	aaatagttata	19260
aataatatta	gcctacatat	ttcttctgtt	ctctttttca	actcttagaa	tcagagtatg	19320
gtagtctcag	actaaaccag	gagcttcccta	tctgtttctc	tgttcttaag	tcactttatat	19380
aatgtaagga	tgctgtgtat	atctgccagc	taggccttat	atacaaaagg	cacctatcac	19440
aaccttctaa	aacagtctta	ccacttagag	accatgttca	aacatatggg	cctttgaggt	19500
aattgccaca	ttcaagctat	aatatgttta	tctaagggaa	tatcttcact	tctagcagat	19560
gcctaaaaat	atctaaagg	aaacactggg	aattgctgtg	tttgttgatg	ctgctcttcc	19620
tcctcctcct	cctcctcttc	ctcctcttct	tcctcctcct	cttcttccct	ctcctgcttc	19680
tccttttctt	catcctcctt	tcttttctta	tttttgaggc	atgatttcac	catgtagccc	19740
taggttaaccc	gtaacttact	atgtatgtag	accaggctag	cctctgtctc	ctgagtgtct	19800
atattaaagg	tgtgtatcac	catatccagc	aacacttgct	ttgagatggg	tagaggaaaa	19860
aaaaatatac	gtaaataaag	atggatgcc	attactaaat	tgttacttcc	agtcaaaact	19920
tgtacctagt	ctaaggccaa	aatagggatt	ttttttctac	tttgcaagtt	ggctccatta	19980
agaggctttt	cttctcttgg	tctcactaga	taggaaggag	agagaggagg	gaaggagaga	20040
aagcgggtga	ggagtgggag	gtagtgtgac	cgagaatacc	cagtaggctc	atatatttta	20100
atatttggtc	cctagtgtgat	agaactgttt	agaaagatta	ggaagcatgt	cttaggggct	20160
ttgaggttcc	aaaatttaat	gctagaccca	gtctttcaag	ggagggggcg	gtctgtctct	20220
ctctgcctgc	tgcatgcaga	gctctcagct	actactctag	tgtcaagcct	gtgtgcttcc	20280
tgcttcaatg	atcataaatt	aactgtaagc	aagcctccaa	ttaaattgctt	tcttttatag	20340
ttaccgtgat	catggtgtct	cttcacagaa	atagtaacct	gtggtgattt	taatatgcct	20400
ggaccaggga	gtggcacttt	taggaggaat	ggccttgtaa	agagggaagt	tgtctctgtg	20460
gggggtggga	atgagaccct	cgtcctaacc	atgtgagaac	cactcttctc	ctattggcct	20520
tcagatgaag	atgtagaact	ctcagatcca	cctgcaccat	gtctgcctgg	aagctgcctt	20580
tgttcccacc	ttgctgcccc	aattaaatgt	tgtacttata	agaattgttt	ttggggggct	20640
ggagagatgg	ctcagcagtt	aagagtactg	actgctcttc	cagaggctct	gagttcaatt	20700
cccagcaacc	acatggtggc	tcacaacccat	ctgtaatggg	atctgatgcc	acctctggt	20760
gtgtcagaag	acaggacagt	ataccacat	acattaaata	aataaataaa	taaaataata	20820
aattcttttt	aaaaaagaat	tgctttggct	atggtgtctg	ttcacagcag	taaaacccta	20880
acataaccct	gactaagaca	acaagtgagg	aaagtggttg	tgtgacactc	tggtatctctg	20940
gaagctcacc	tcagcatgaa	gcttgctcga	gcgnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21060
nnnnnnnnnn	nnnatctaag	tacactgtac	tgcttctcaga	cacaccagaa	gagggtgtca	21120
gatctcatga	cagaggttgt	gaactcagac	ctttggaaaga	gcaatcagtg	ctcttaactg	21180
ctgagcatct	ctccagccca	aaataattct	tactagtaac	atggaacaat	caagttttat	21240
tatatgatac	atattaatca	acttataagt	acatgattat	gcacatttat	catatcgtgc	21300
aaccatcact	gctgtcgttt	tgttttgttt	tgttcttttg	aggcccggtt	tctgtgttgt	21360
tctggaactc	actctgtaga	ccaggctggg	cttgaactca	atgatctgcc	tgctctgcc	21420
tcccaagtgc	tgaaaacaaa	tgtgtgcacc	accacctctg	gctatcactg	ctgtcttttt	21480

tttttttttta	acagttattt	atttcgtgca	tgcattgtgtg	tataagcatg	taacgtatgc	21540
catggtatgc	atgttgaggt	cagaggacaa	ctttcaggag	ttagttcttt	cctcccactg	21600
tgggttctag	gaaccaagct	caggttggtt	gacttgcattg	gcaagtgcct	ttaccacaga	21660
gccatcctgc	tggccctact	ataggtcctt	atataaaaaa	atcatatgcc	gggcaaaaac	21720
caaacaaaaa	ataaacctca	aaaaacaaaa	ggaccatata	atattgtggg	ggagtgagtg	21780
aagtccctgaa	cgaatgtgtt	ctgttgacat	gtctgtactt	cagacccatg	ggaattggca	21840
aagccttcct	ctggctctgt	gaggatgctg	atagttctgtc	taaaaaactag	agatcacagc	21900
tttctcctct	ggatgactgt	aaccccgat	tgttctctct	cagagactgt	ccaccaagct	21960
accctgccta	cttaagctgt	acacaatgaa	tgagctgagt	ttccagggtta	cagcacagta	22020
gacactgtcc	atcagtgaga	gcacagccta	gcctaacagt	acacatgtct	gctttcttca	22080
cgtttccaga	accaagcctt	gctggataga	gcataattgt	ctgtttggct	tatttcaact	22140
gataaaaagt	tttcaaggag	ggccagggtg	ggtggcacac	gccttttagtc	ccagcactcg	22200
ggaggcagag	gcaggcaaat	ttctgagttc	gatgccagcc	tggtctacaa	agtgagttcc	22260
aggacagcca	gggctatata	gagaaaccc	gtctcaaaaa	acaaaaaaaa	acaaaaacaa	22320
aaacaaacaa	caaacaaaca	aaaagccaaa	aatccaaccc	cccccaaaaa	aaaaacacaa	22380
ccaaaaacca	aaaaacaaca	acaacaaaaa	gtttttgagg	tttaatttat	tgcatgtcac	22440
agaattttcac	tgttttaaaaa	aatggctgaa	taatatattca	ctatccattc	acgtatttgt	22500
aggcattcat	gtgtgttagtg	gttttaataa	aaatagcccc	cataggcttc	tacagttgaa	22560
tgcttagtca	ttgagtagca	gtactagaga	gggaattgaa	ggtgtggcct	tattggagta	22620
ggagtggcct	tgttgacagga	attgtgtcac	tttgagggtcc	cagcaacaag	gttgctctga	22680
tcacatccaa	agacattcta	ggtctatgtg	atctggctgg	aattcagaca	tgcccttaat	22740
acacaccttt	aatcccaaac	aatgaaggta	aagttagttt	ataaaaaaga	gcacccatgt	22800
ttgaaagtga	cgttttaatta	agagtgatga	attagagaaa	gatctgctgt	cacagagcag	22860
agaggaaaga	gaggcagcat	aagagggagc	atggcagagg	gagagggagg	aggggttttc	22920
accagggcat	ttgtacagag	acaggttgca	gagctagaac	aggtgaagac	agaacaagcc	22980
agagaatgag	aaggagccag	gagattagga	cagattgcca	atgttaatag	gctaagcaga	23040
gcatttttagt	cagaaactga	gagaagtcaa	attgaatcag	ttagcttgga	aaggagtttg	23100
agcagcaaca	gctgagttaa	actagccaac	agaatccaga	aagaactaga	aaagatgagc	23160
ttactcagca	gcaaatctca	gaggctaaaa	acatcttaga	cctagattag	actgcatgga	23220
ggctagacgc	ttccagggct	aggcctaggt	tagcagacgg	agagagtaat	aagccttgga	23280
gacaacagtt	aatacagaag	actatgtaca	gacatggata	tgaacctctc	agccacttct	23340
ccagcgtcat	gcctgtctgc	attgttagga	gtcatctagg	aaaggctaag	ggcaggcaag	23400
caacttttcc	agagatggct	caactgtttt	tgcatggcct	ttgagaggcg	agctctgaga	23460
gggaagggtc	caagagactt	catcccagga	ttgtctgctta	attacgacat	gccttttctt	23520
gtcactgtta	tttagtataa	tgactcctga	gctttagccc	atcctatttg	gcataatttc	23580
tgcagatcaa	cataaagatg	aactttcaca	aatataatgct	gttttagatga	ataaatgatt	23640
ttataaaatt	cctgatttga	tttaataaat	tttaggaaga	aagctttagg	agatagttta	23700
gttggtttgc	cagaaaagatg	taataacgct	agaatcaaga	atagaatgtg	gtcgggcagt	23760
gggtggcagat	gccttttaac	ctagcacttc	ggaggcagag	ataggcggat	ttctgagttc	23820
gaggacagcc	tggtctacag	agtgagttcc	aggacagcca	gggctacaca	gagaaaccct	23880
gtcttgaaaa	acaaaacaaa	aagaaaagta	agtaaaaggct	gcataataaa	gaataacaatg	23940
agcttttcaca	actacaccaa	aaagagacat	gcttgggaca	aattttgtgat	caaggaaaaa	24000
tattcatttc	agatcaggct	caaggatgaa	gccacaagtg	tgtgatatga	tgaacaagac	24060
catggataaa	ctgttggttt	gagcttaaa	aataaaacac	tgctttgaaa	tttaactatca	24120
acattctact	gtaacttttc	tttttataaa	ttttatctat	gagataaatt	tctaaagaac	24180
ttgtgtctat	aaaggtatag	aaggacagag	agaaagaaat	aaggtgtggc	atctgggctc	24240
tgctccatcc	acccaaataa	atatgtgtgt	gtgtgtatgt	atgtatgtat	gtttatctat	24300
atgtatgtat	atacatatcat	gtgtaggtag	gtatatgtgt	atgtatatata	gtatgcatga	24360
acacttgga	agttgatgag	acaagtgaga	gggtgggcc	ccnnnnnnnn	nnnnnnnnnn	24420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	25320

tattcatgct	tgggtaactg	gccctggctg	attgatgaga	gctgatttcc	ccactgcctt	25380
gtggcagggg	ccactgcgcc	cacagggctc	cctcaggatc	ctcagtagag	agctgcacag	25440
ctgggtggaa	gtagagggtc	gcatatataa	cacgatctca	actttatttc	tttaaataaa	25500
aatTTTTttt	aaatTTTTta	cagctctata	taaacgaagg	aactattgaa	gggtcagcaa	25560
ggacctgcca	acggttgtca	agggtaatgg	cgatgtagtg	attttttttc	ccccctccat	25620
tttacttcca	tactttctac	attaccccac	aactggcgaag	tattatttta	aaatgaaagt	25680
aaatagtga	agatgacttt	gaaggaaaat	tgaatcggta	aaaagaaagc	tgagagacca	25740
cccgggaagc	ccaggctaaa	tgtaatctgg	gtcaggccctc	ccaggccctgg	gggtctcaaga	25800
tggtcagctg	agggaccctg	gtgaccctct	tgggccagca	gggacgggga	ggagccggaa	25860
gctgagtacc	caaagtgtct	ctctgggctc	caagggcctg	cacagagact	gtgtgggaat	25920
caaaggatac	aggcatgagg	actgaggcct	gacgaaccca	gctatcattc	gtcctagaac	25980
aggaggcaga	gctccaagag	tccaaccaag	aggcaggaag	ctttgggacc	cgagatgggc	26040
gatgggatta	gaaaggcatg	tttgcaaata	ctttcaaatt	tacgatgcac	actcactgga	26100
aacccccacc	ctgggtgtcc	cttccctgcc	tcttgccaca	cccaatagct	gacatcactg	26160
gagaaagtcc	caagaccagg	ctggctggag	ctcctgatag	gttccaccct	cctgcagagg	26220
gccctcgaag	actagcttgc	tcgcccacac	cgccagatgt	ctgtgtcttt	ctctcttttg	26280
cctccccacc	tcgtctcttc	ctccaacctc	agtgagggtt	cccctgcttc	ctggggaaag	26340
tagaacttgc	cagtgtcac	tgtaatgtcg	tccctgtagg	tgtcatggtc	ccccattact	26400
gggagcaggt	atgcctcaga	tctccctcta	tctcgtgccc	tttcaggctg	tctcagtttc	26460
tctctgacag	ttcctctcct	cctgaatcct	gcttggtggc	atgcgaacag	gctcaatata	26520
ttccatctca	aaaaacaaac	actgggaagg	tgttgagaga	cagagagcat	gggtaatggg	26580
tgccccagct	tggctgggaa	ggggtaactt	acaatgctct	actgccaggt	agggtagctg	26640
cagttgtcaa	ttaatgttaa	atttcaaaat	agctagtaga	gaggatttta	gatgttccca	26700
atcccaacac	aaagaaatga	taaacattca	aggcgatggg	tatgctaatt	gctctgatct	26760
gatcacccga	cattgtatac	atgtttttga	aatgtcaggc	tgtaacccat	aaatatgtac	26820
aattaccgtg	cagtgtattca	agataaaaaa	tataatttta	aaaagctaaa	aacagaaggga	26880
aatagctgcc	cttgaccccc	ccacccccac	aaggctcctc	ctgtttgtcc	agccacttaa	26940
tgtcagagct	tcctgtggga	gggtgggttt	gggtgacaca	gacactcctt	cctccctcct	27000
tcccataaag	aggagtcaac	cctgtcccac	gatgccatgc	agggccacat	gcgtgatatt	27060
aaccagtaag	atgtgagcag	ggatgatacc	tgtctcttat	aacaaacgga	aaaaaaaaacca	27120
caccaaacca	aaaacaaaca	aaacaaacaa	caaacaaaaa	cagggttggg	ctgtccctgt	27180
gtcttttccc	acataaaagt	aagcacacaa	agtagccacc	atttatttat	ttgtcccttc	27240
ccccaccctt	ccccgagaca	atgtttctct	gtataacagc	cctagctgtc	ttggaactca	27300
ttttgtagac	caggctggcc	tggaaactcac	agagacacag	agattcacct	gcctctgcct	27360
cccaaagtca	gggattaaaa	gcatgagcca	cgaactaacc	agtaccccag	agctcttgac	27420
tctagctgca	tacgtatcaa	aagatgacct	agttggccat	cactggaaaag	agaggcccat	27480
tggacacgca	aactgtatat	gcctcagtac	aggggaacgc	caggggccaaa	aaaatgggaa	27540
tgggtgggta	gggaagtggg	ggggagggta	tggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27660
nnnnnnnnnn	nnnggcagga	tcctgtgttc	atgtgcaaca	ctcgatgcaa	gctgtgtagt	27720
gtttgggtct	gagcacctga	aggggaccaa	gcaggctgat	gccaggcca	cgggttcttt	27780
ctggccccac	tgccccactc	caccctctgg	catccccatg	atgaacatgg	ccacagatca	27840
caactactctg	ctcctctccc	agatccacgg	agccataggg	tcccagatt	catctctgca	27900
gctaacaagc	tgggcagtgt	cacctccctc	aaggttcctt	tcctgtctctg	agcagcagtg	27960
tctcccacag	tgagacactc	atgtccactg	gaagatattg	tagccattaa	attcctgtgc	28020
taaaataaact	agggggactt	gtcaatcact	acactcttag	ccccggactt	ctgactcata	28080
gaggggtggg	acagctcagg	gacctgcatt	ctaccaaaata	gccatgtgtc	cctgatggag	28140
gaactgcccc	tggacaaacct	ctgcagcaac	tgaacctctc	gtgggtctcct	agttcttctg	28200
gacaggtgtg	accccagtac	ctagtgccag	gtgagagagt	gctagggccca	cactaagggg	28260
tgacaggaca	aggttggagc	tggtagatgt	ttggggccacc	aaagagaaca	ggtcagtagt	28320
aaaagccatc	atggcctgag	ccagcctgcg	agtctcctct	gcagttggga	cactcttgca	28380
gtgtcctggg	gacctcttga	gggtagcatg	gtcaccaaaa	tcctacaagg	acagatcaga	28440
agtcagttag	gtcaagggaa	cagctctagg	ttctctgtgt	ccctcacgga	cccttttttt	28500
tttttttttt	tttttaagat	ttattttatt	attatatgta	agtaactgtg	agctgtcttc	28560
agacagctcc	agaagagggc	atcagatttc	gttacggatg	gttgtgagcc	accatgtggg	28620
tgctgggatt	tgaactcagg	accttcggaa	gagcagtcgg	tgctcttaac	cactgagcca	28680
tctctccagc	cccctctcag	tcctgatgcg	acagggcagc	aaaggccttg	tcccagatct	28740
gaggagagtc	atgctgaagt	ccttcctacc	ccaccccttc	cgaacccctg	aacatcagcc	28800
ccataactac	tgactcccc	acccccattc	ccttgcttcc	actgatccgg	tcctcctctt	28860
ccctctggcc	ccaccattc	ttccccagcc	ccacctgatt	gtacctgggt	gtccaacttg	28920
aagagggcag	gcaggggcag	cttctgctgg	gcctgctcac	tcactggctg	tagaaatgag	28980
aaaggagatg	aagaaaaggc	ccttcccatg	ggtccccatc	ttgccaagac	ataggtgagt	29040
ccctttggct	cttcccccta	aacctctcac	ttttgagtac	ctgctggccc	gggagatcca	29100
cggcgctcac	cggagagaac	tgttgagaaa	agggagaaca	gagaactcag	cgttcctccc	29160

tctccaccct	tctggcctct	cccagatttg	ccccgcccc	ccagcatctc	cttcagcctg	29220
actgaccact	ttccactcag	acctcagctc	tgcctcaccg	tgaaacaggg	accttgacgg	29280
caggacaagc	tgagtacgag	gagcccccg	agcagtgcc	tgctcctgta	tccagaacag	29340
ggagtgttag	ttcctacctc	acgctcgaag	gccaaagcag	agactgctat	ccatggggtc	29400
cttgaccgca	ccaggetgcg	gaacctggac	tcaaaacata	gcagctgtgg	acctcactca	29460
ctctgagagg	tgggatttcc	ataagctttt	tttttcacct	gtacatttag	tcttcattct	29520
tttcgtctta	cactgtggat	cagtcctggg	ttcaaattta	aagccctcat	cttgcaagag	29580
gacctgtgcg	atctcccttc	atgcctttgg	ctttaccctg	tcttggtaat	tcatggcaga	29640
agttcttctc	gctcccctgt	agatgttgag	gacccaaata	agaatctctg	taaatactga	29700
gcatgatgcc	tggccccccac	cctagcaaag	ccacctgacc	tggtgttcat	ttcatccagc	29760
ctttctcagg	ctgccttggt	cctaccctaa	ggctctgaga	gctaactctg	gctggcaggg	29820
cagccagaaa	cttctttgtt	gaccaatgaa	tgactggccc	agacaccttt	ggacttacgg	29880
gaactacaag	cctcatccca	cttctgtctc	aagttctgat	ccagggtgct	tcggggaagc	29940
ccagctggcg	gaagggggga	ggctctcagc	ctagagagcc	ttcctttcca	tcctcagccc	30000
cctaccacag	ccttatttca	ggcaccagct	cttctaaaag	gtccttctgt	tatccctaga	30060
cctccacaac	tgtgttcaag	aaccttcagc	cagggcctca	tctccaatct	ggatatatga	30120
tttttctcgc	caagagttag	cctccaggtt	ttggagtctc	agagggttct	cctggagctg	30180
cctggacctc	tgctcctcac	cacccagga	cgctgtgaag	ctgcaggctc	cctgaataaa	30240
ttcatccaga	ccccctgcca	aggtgccagc	tgtctacttc	ctctgctgcc	caagcagcag	30300
gctgcaccac	ccctccatcc	tacctcttca	ggcttcttag	cgagcacac	gcagcacacg	30360
gtgttctcct	ggaccagctt	gctccccacg	ctccccagct	gcagccagca	gggcctagct	30420
ctctcctccc	acaggacctt	tgctctcagc	caacccccgt	tcagcttggt	ttcagtgtctg	30480
gtaaatattg	acctgtacat	ccggttaaac	attgatattg	gggcccagaag	accctttccc	30540
atcaaggcta	cccagaaccc	tgcttgagcc	tggagaaggg	gtttacagga	gcagataagt	30600
gaggaggttg	ggcctggcaa	gccttctaat	gatccctcaa	cataggggat	tatccacagt	30660
cagtgaggct	cagagaggct	gtgtggcctg	tgtaagggcg	cagagtgggc	tcagagtcca	30720
cagccaaagt	cccaccacca	ccaccaccac	caccaccacc	accaccacca	ccaccaccac	30780
caccaccacc	accaccacca	ccaccaccac	caccaccact	accaccacca	ccaccaccac	30840
caccaccacc	accaccacca	ccaccaccac	cacctcatct	accatacta	anttgaggct	30900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	30960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31320
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	31980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	32940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	33000

ggctgaactg	atgggcagct	aaggtccaga	cagtggctgg	ctcttgga	gcctgtctct	33060
ttcctttgac	tcagaccact	ccctgccgtg	gcttacatca	ggaggtgcaa	gggctgcagg	33120
agggcagcca	gacccacaaa	accagctagg	ctaaatggtg	cttattgttc	gcaagaggcc	33180
atgacctcat	ttgtctccca	gctcttttgg	taagagagaa	tgagaggaag	ctggacagag	33240
aacctagcag	gcctcaggca	gcccactgct	ccttgctgta	aggaaccag	caccgatggg	33300
tctgaaaagc	agcgatccga	atggagtcag	gctgagctgc	aggaagctca	ccttccttgc	33360
tcactgctgg	tggaagcaac	ttcaggaaga	gcccagccta	tgggactata	gctcctccgg	33420
ggtactgctg	agtccagccc	cagagcttag	ctccctgctt	cccaccaccc	accaccacat	33480
cctttcccaa	caccattcaa	aaccccagtc	cagcctctcc	tactggtcta	cagtgcagg	33540
ctaataagag	cctgggcctc	tgtcccccca	attctctctc	ccctctcatc	tgttcacctt	33600
ggttcctaaa	ctgcaggggc	tactataaac	ctacctccac	ttccttgca	ccctcttttc	33660
tgctctctgg	ggtgcccctg	ccactcccag	tccctctagc	cagggagcct	cttccatata	33720
tgtcttcccc	aggctagacc	aggcgtgccc	ttacctgtgg	ttgcggcagc	ttctctcaca	33780
gcctgcactc	tgaggggctc	caggaagcag	tgaggggagt	agctgcctct	caaccagcgt	33840
ccagcaggct	tcagattaca	gctactcttt	tcttaaagtg	acctgactcc	atttggaaatc	33900
tgtgattgca	tcattgtctg	gtgttaactt	taaccactg	ctgcccttcc	gccatgtggc	33960
tccaagacca	cacgttgccc	accctcctct	cccaccacat	ctcccttgga	tctttatctc	34020
tcttcattgg	gaccttcatt	gggacatgat	ggctaacttc	aggggcactt	gggcccagcct	34080
ggggtaggtc	atgagtctga	acttgaacat	ctgaaaggat	tggtctgagag	gcaggctgca	34140
tgagagagact	gtgagccagc	cgtatgggag	atgctggggt	cttccaggcg	cttggctctg	34200
gctcactgca	ggtgggagca	agggtattct	tctccccctc	tcacctggaa	aatgaaggaa	34260
tgggactgta	cctgacagct	ctgaaggctc	caaaggacag	tggggtgggg	actagagagt	34320
tggcccagtg	cttatgagca	ctggctgctc	tcgcagagga	cctgagttct	gttcccagct	34380
cacatagcaa	ggactcgaaa	ctgcttgga	ctccagctcc	agagaatctg	acgctatctg	34440
ctnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34560
catgccttta	atcctagaac	tcaggaggca	aagcagatgg	acatctgaat	ctggtggtg	34620
cattgtctac	atagcaagtt	ctaggctagc	caggtacatc	ataagaacct	ttctcaaaaa	34680
ataaataggg	ccagttggca	aaatttagct	tgccctccta	acacaagaac	ccaaagtcaa	34740
ccccagcaac	catgtaaaaa	gaagcaaggt	gtggtggcac	ttgcttgtga	tccagcattg	34800
tcaaggtgga	gacagacgga	tccatggggc	tcactggcca	gccagctagc	tggtctactt	34860
agtatgctcc	cagccagtga	gagactgaaa	aataaataaa	taaataaggg	gttaggaaga	34920
ggtaacatgg	tggtcagtg	agaaaagata	cttgtcatal	aagcctagca	accctcaatt	34980
ttcagtggcc	actaaagggt	gaaggagaga	accaactcca	aagaattgtc	tcctgacagt	35040
tttatgctgt	ggtacacaca	cacgcacaca	cacacttggt	acatgcatat	gcatacaatt	35100
aataatttaa	aatgttatgt	gtatgggtgt	tttgccctaca	tgcatatctt	tctgtgcacc	35160
acatgtgtgc	aatgcctgtg	aaggctagaa	gaggacatca	gatccccctg	gagttacaca	35220
gggttggttag	ctaccatgtg	gattctggga	acaaaacat	gggttttcca	aaagaggctc	35280
ttaatcactg	agccatctct	ccatccccct	aatagtatat	atttctgggg	ctggagagat	35340
ggctcagagg	ttaagagtat	tgactgctct	tccagagggt	ctgagttcaa	ttcccagcaa	35400
ccacatgggt	gctcacaacc	atctgtaatg	gaatctgatg	ccctcttctg	ctgtgtctga	35460
agacagctgc	agtgtactca	tatacataaa	gtaaaataat	aaaccttttt	tttttttgtt	35520
ttgttttttt	tggtttttcga	gacagggaga	cagggtttct	ctgtatagcc	ctggctgtcc	35580
tgaactcgc	tctgtagacc	aggctggcct	tgaactcaga	aatccgcctg	cctctgcctc	35640
ccaagtgtctg	ggattcaggg	tttgcgccac	cgccaccacc	tccaaggctg	ctgctgcggc	35700
caccaccacc	ccaccaccac	actacctgac	tatttaactt	ttaaaggcag	ccatctcatg	35760
gaaaatgaca	cctagcattg	tcctctggtc	cctacatgac	cccatgtgca	aacacatacc	35820
tgcataaaca	cacataaata	cataagtaaa	cttagtctgg	ttgttttgga	aatgtgctat	35880
ggtttggatt	gtgtccccc	aagggaaaaa	ttgggtccca	gagtagtact	attggaagag	35940
agtagaactg	ttagggttta	ggcctggtgg	gaagtggcca	tggttagaga	gacatgccaa	36000
ccaaggggaa	tctctggctt	catcttttcc	ctttgctttc	aggctcctaag	acagtcaaaa	36060
ggctgcttca	ccacatgccc	agaagcaagg	ggagcagtca	tggtgggac	cgctaattgca	36120
gctgttgatt	tccccagata	tttgtagttag	taagagacag	gtgaggaacc	ccacagcaag	36180
tgttagtaat	tgtgtgtgga	ggtgccctcc	ggggacgggg	gccctcctgg	ggcaggacgt	36240
tcctcttcct	catccacctg	cactccgaga	acaggaaatg	gtgactttgg	cagagcttaa	36300
gcagagcccc	ttcatgttac	aagtatgtaa	attcataagg	accagtttct	ctccatatga	36360
aacagcttca	aacaggagaa	ggaagaagca	aacattaagg	aaaagctctt	ttattgcaga	36420
ggctacactg	aagctaccgg	ccgccttctt	ggaatgtata	atcagcttcc	ctctgggggt	36480
tctgtagagc	actgagacat	taagtactac	tggggtccag	gattctgcct	atgaagagga	36540
gggcccccg	gtccgtgtcc	ctcagaacaa	agaggaaagg	ttggttaagg	tgatagtcta	36600
gcgggaagg	gagggcgacg	ggctggaggc	ctgggctggg	gctgcttctt	gccccctctt	36660
cattccactc	gaaagcagcc	ctgtgttcca	cttgggtgag	cttcacgggt	ttgccagtaa	36720
tcttgctgaa	gtcgggtgat	tcaaacaacg	actgtagctc	tgtggagatt	cagagattcc	36780
attaacacca	cacacacaca	cacacacaca	cacacacaca	cactccctgt	ttgtgtaggc	36840

tgatttttcaa	gaaagcaagc	tagaagtggg	gtacctcaca	gtgacttggtg	agctatgagg	36900
cactctgtga	caggctcagt	gacctacctg	agaacttata	gccaatggtg	ctgaagccag	36960
acctggcctg	agagaatggt	ttgggctgtt	ataggacaca	tagagataca	cacacacaca	37020
acacacacac	acaccaagga	ctgagtctaa	tgggaggtgg	ttcttcatte	ccctcccctg	37080
taatggtgtc	acatgttccc	tgagccaccc	tacaaagaaa	gccacaggac	tcagttctgt	37140
cagcaaggtg	gcaggctcca	agactcagcc	ccgagcgcaa	agtggccttg	caaacatact	37200
catgtctctg	agagacttgg	taagtctgcc	ttcgaagctc	agcttcagct	tggggacagt	37260
cagcacagct	tggatagtct	tcagttctcg	gtcgaatgtc	tgaatgaact	cagaggtgag	37320
gctctcttct	atcatggtca	agttctgggt	cacggtcagg	ggcaggaaga	agatgatgct	37380
catacttcc	gtcaagggca	gctgggcaat	ctaaccacac	agagatgcgc	acaggttagt	37440
tgtgagccag	aaaaaacaaa	acaaacaaac	aaaaaacac	caacagctgc	cttcccctct	37500
gctgtaacgg	ggccccagcc	ttgtgctccc	cagcctcagc	ctgggctgta	ggctactggt	37560
tactggcagt	ccttccatga	gtagggagtt	ttcttctcag	cctaaaaccc	acagaagttt	37620
aatgaacaca	cgtttgtttg	tggttccgct	acggtttcta	ttgtgataaa	acatgactga	37680
aagcaacttg	gagaggaaag	ggtttatttc	atctgacaat	tcgcagggtg	tcttctcatc	37740
actaagggga	ctcagggcag	gaactgaagc	ggaagccgtg	gaggaacgct	gctttctggc	37800
ttgctccccg	tggcttctta	gcctgctttt	ttatgctatc	cagaaccact	tgccccaggag	37860
tgacactgcc	cattgtgggg	tggggccccc	cacatcaatc	actaatctag	aaaatgaccc	37920
acgggtttgc	ccagaggcca	gtctggtggg	ggcattttat	caattgagtt	tcacccttcc	37980
aatgactct	aacttgtgtc	aagttgacca	cacgaatcag	ggcctggttc	ttaggagctg	38040
aagtggaaatg	ccccccagag	actgcctgcc	agcactgctg	accatttgct	ttgtatagag	38100
cattggaacca	gaaatgaaca	ataaaatgga	tcctttgaac	agatgtgttg	atcctagggc	38160
ctgtggacac	agcgactggg	cttcccagag	cccccatgga	atcannnnnn	nnnnnnnnnn	38220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38340
gagagagaga	gagagagaga	ggataaaagg	ttagcccagt	ggtggtggca	cataccttta	38400
attccaacac	ttgagaggca	gaggcagggg	gagctctgtg	ggccagtttg	gtttacagag	38460
taagtttcag	aatagccagg	gctacacaga	gaaacccctg	cttgaagaga	aacacacaca	38520
cacacacaca	cacacacaca	cacacacaaa	taagatcttt	aagaagaaaa	gaaaggatag	38580
tggggaaaca	tctgagcaga	ggaagaaatg	gggtgcgcag	gacacccacc	ctcagaggag	38640
gccctcactg	gaggtgtctg	cacaggagaa	cacttgcaat	cagcttgccc	tagggcgctca	38700
gaactcagaa	ttcagtttca	aagcactgac	aggagcagtg	actggggacc	ccaggttgaa	38760
tcccccttt	atctaaaatg	agtaagaacc	aagtgttttg	aagtggttgg	gatttggaat	38820
ctgggttatt	tgcatctaca	aaagaggtct	tggggaggaa	acccagttct	acccctggaa	38880
ttcatcagtt	tcctataccg	ctgactacac	aggggctgaa	ggtaatctca	atgttttcat	38940
aactggtgtg	gtgctacttg	ctcatgatcc	caacacttgg	aaggtaggtc	agaagttcaa	39000
gagcagctct	gactactcag	tgaatttgag	gctagcctgg	gctacatgaa	aactcataaa	39060
acaataaaag	aaaagaaaaa	gtggcgagtg	aggtggccca	tcaggtaaaa	gtgccaacca	39120
cctgcctga	aagcctccac	acggaagggg	aaagccagct	tctacatgtg	gtcctctgcc	39180
ctccgcatgc	accatggctc	gtgcaccccc	acacccaccc	acccacccac	ccacatgaca	39240
taaataacttg	taatgattag	tttctgaaga	acaatatatt	cgttgatctt	gtttagggaa	39300
caaagtctgt	gcacattgac	ctgtcgaccc	tgtagtacgg	gatccgctcc	aggaagctaa	39360
agattttggg	atgcttcatg	ttgcagctac	tctgatgaac	agtgttgctt	tctgggccag	39420
gtaattggtg	catatacctt	tgatcccagc	actttggagg	cagaagcatg	tagatctctg	39480
tgagttcgag	atcagcctgg	tctacagagt	gagttccagg	atatccaagg	ctatacagaa	39540
aaacccctgt	ctctaaaaat	cactaattta	aaaaaaaatt	cctttctaaa	cctatataac	39600
aaatgttttg	taggctgcct	taacaaaagc	caatggccat	tcagagaagg	ctcaaaaagag	39660
aaagtttagg	ggactctaca	agcatcctca	ggaaggccac	agaaaagcaga	gcctgggcca	39720
gtgagacttt	gcagtgggca	aggttcagct	ctttatgtag	gaagaagaga	gtcaacagtc	39780
agagtccagc	tttccataaa	acctgtgcag	ggcctctagg	caaagccctg	tgttaggggc	39840
aaaggcattt	gcagtctaag	cccgttgaca	tgagctcaat	ccttggaacc	caggtggaag	39900
gagtgtgctg	actccacaaa	tttgtcctct	gatctataca	tgtatgcacg	tgcaacgcaca	39960
ctcacatata	tgtccacatg	cacacatgca	tatacatgcg	catgctgtga	cccacacaca	40020
gggtcaaaaag	cagcaagaga	tgccctgtga	aaaacgtctc	attcagctctc	ccatcatcca	40080
gtgccacact	ctgagcacag	gtggtactga	tatcgcttct	gattgatcga	tcagttgatt	40140
tgagaccccc	cctcactatg	tagcccaggc	tggcctggaa	ctcacagtg	tcctcttgct	40200
tctgcaagat	gagcccatca	tgcccaccat	gttattgaag	caataccatg	ctctataaaag	40260
caaacctagg	caggcaggat	ggtggactcc	tgtaatctca	ggacttgaaa	agtagaagg	40320
agatgaggag	ttcacatcaa	tctcccgtat	gcgttggagg	ctggagtgga	tgttccctgg	40380
gcgcttctgc	cagcacctga	ccaatgcaga	tgcagatgct	cacagccaac	catcagactg	40440
agctcgggac	cccagtgagg	gtgctggggg	gaggactgga	ggagctgaga	cgggattgca	40500
agcccatagg	aagaacaatg	tcagctggcc	aaaccaccca	gagctcccag	ggactagacc	40560
acgaaccgag	gactgcacat	gaagggatcc	atggctccag	atgcatatgc	agcagaggac	40620
agccttgtct	gacagcatgg	gaggggaggg	cattggtcct	gtggaggttt	gatgccccag	40680

tggttgagga	tgctggagcg	gtggggcagg	agtgggtggg	taggtgggga	gcaccttcat	40740
agaggcaaaa	gggatggggg	agaaggcaga	tgggatgggg	gggttggtga	ggggtaagaa	40800
agaaaaaaga	tgctctctgaa	agtaaaaagt	acttgctact	aagcatgagg	atatgagtca	40860
acccccaaagc	cccacagggt	ggaaggagag	aaatgagtcc	cacaagttat	tttctgatct	40920
atagctgcaa	tccatggcat	acgcagaaac	gcaaagacag	acaatgagtt	gggtgtggtg	40980
gtgcacatgt	aattccatca	ttcaggagac	agaagcagca	gagttgttgg	aaatctaagg	41040
ccaacctaaa	gacctacacc	caataaaaaat	caaaactataa	ggaaaaaggt	ggtcgaccaa	41100
tgtaacatta	aagttagaaa	tctctcttca	caactgtgtag	atactgtaca	aggaagagaa	41160
aaggcagcca	catcaaaaaca	gtgtaaatca	acgagaaaac	cagaaaacaac	tcaagagaag	41220
gctgcagggg	cctgaattct	gttctcagaa	cctgcatcaa	gccaaagaaa	tcaaaaactgt	41280
ctgtaactcc	agctccctgg	gatccaacac	ccatttctgg	cctccatcag	catcactcac	41340
agggtgtgac	acatacacat	caataaaaaat	caaaaccagg	gatgaagggg	tagggagggtg	41400
catgtgtgac	tgaggaggagc	tgagaggcac	tggtgtaata	caataaaaaa	tttgggtgcat	41460
gggtggtgcac	gcctttaatc	ccagcacttg	ggaggcagag	gcaggcgaat	ttctgagttc	41520
aagaccagcc	tggtctacag	agttagttcc	aggacagcca	ggtctacaca	gagaaaccct	41580
gtctcaaaaa	aacaaaacag	ccggggcgggtg	gtggcacacg	cctttaatcc	cagcacttgg	41640
gaggcagagg	caggtggatt	tctgagttcg	aggccagcct	ggtctacaaa	gtgagttcca	41700
ggacagccag	ggctacacag	agaaaccctg	tcttgaaata	aataagcatt	tggtgctggt	41760
acagaaaact	ccagcccagt	ttccagcaca	cacagggtga	ctcacaacat	cataactcca	41820
cttccagggg	atccaatgcc	ttcttctgac	ctctgtgggc	accaggattg	catacagtgc	41880
acagacatgc	acataggcaa	aacactcaca	aaataaaaata	aatctagcaa	aaaaaatttt	41940
aactaataat	ttaaagaaaa	aaataaggaa	gccgggggtg	gtgtcgcacg	cctttaatcc	42000
tagcacttgg	gagacagagg	caggcggatt	tctgagttcg	aggccagcct	ggtctacaaa	42060
agtgagttcc	aggacagcca	gggctacaca	gagaaaccct	gtcttgaaat	aaataaataa	42120
ataaaaaata	aggccaagta	attcttggaa	gaatcccaag	gggacactaa	gtgtatataa	42180
aggcggtcca	tagggctagg	aatgaggctt	agcgagagca	acttcgctgg	tgatgaaag	42240
tccctcagct	gcatgtggtg	cctttaatct	aggctctccc	gaagcagagg	cagaaggatt	42300
tctgtgagtt	caaggccagc	ctggtgtaca	tagctagttc	caggacagaa	aggcgatat	42360
aatagaaca	tctacctag	agccngcca	anaaaagggg	agacctgaga	ccagagagat	42420
gactcagtg	ctaagagcat	tgactgctct	tccagaagtc	ntgagttcaa	ttcccagcta	42480
aaaattttatt	taaatgttta	ttacttgtat	tattatttaa	atttaaataa	ataagtaaat	42540
gggagcctag	gtttgagtc	ccaaatcacc	aagaaaaaat	gttatcattg	ctaataatca	42600
aattaagagc	ataagaactt	ctttttaaag	aattcttatt	tattttatgt	attgtaagac	42660
actgtagctg	tcttcagaca	caccagaaga	gggcattgga	tcccattaca	gatgggtgtg	42720
agccaccatg	tagttgctgg	gaattgacct	caggacctct	agaagagcag	tctgtgctct	42780
taagtactga	gccatctcta	cagctcttat	cagggtgata	aaatttaatc	tcgtggagcg	42840
ctgagaccaa	gaactaaagc	tgggagattg	aaaaatgcag	accaccaagg	ccctgctcat	42900
ttctccagtt	ctgatcagct	cccgtaccag	gggtctaacc	aggcctgtgt	ctgcttccct	42960
gagtagacca	gaggccccat	ctaaacagcc	tgccctgcagc	agctcctctc	tctaggtgga	43020
cagatgggaa	tttcagacca	atgtcatttc	ccaggacatc	aacacagcag	ccaaatttat	43080
tggtgctgtg	gctgccacag	ttggtgtggc	aggatcaggg	gctggcattg	gcacagtgtc	43140
tgattattgg	ctatgccagg	aaccagttct	tcaagcagca	gctcttctcc	tatgccatgc	43200
tgggggttgc	cctgtctgag	gccatgggac	tcttctgttt	gatgggtcgc	ttcctcatcc	43260
tcttcgcat	gtgaggctcc	ctggggtcac	ccagccgtcc	ctgctgcctt	gactccatgc	43320
cagtcctggt	gctggagtct	actgagattt	accattaaac	agcaacgttt	ctctaaaata	43380
ctattaatta	attaattaat	cacgtgacaa	ccccagcgtc	catatgggtg	tggaatatga	43440
ggaactctac	ccatcataca	tggcgactat	gaagaacaat	gtgacagaaa	atgctaacat	43500
catgtgtgac	cgcattgcac	agccctgact	gctaaaagtg	gacaagcccg	aagcgaaagc	43560
ccaatgttct	acttctaaat	gcatgcacca	aacgcctccc	acaggaccag	agggtcagct	43620
ctgatagggt	ccttgccctg	catgcatgaa	gctgtggaca	cgaggcatta	ttcgcaagaa	43680
cattctagct	gtctggaggg	ccctcaatcc	actgtgttcg	cgctgttcca	gcaccagtgc	43740
ctcctggggc	tgcacctgaa	aaaggggact	gcttaagagg	gctcctacca	agcctactgc	43800
cacagatgca	tgatgggaaa	gccttcttga	agcaactggc	tgccaaaggc	tctggacaag	43860
agatcacctg	ctactggaaa	ggtgggttca	gtctaggttc	tgtgggattc	caggaaatta	43920
gacaacactg	gcagtccaac	agacagacga	tctaaacttc	caaggcacag	ctggtagaac	43980
ttgtctcgga	accagacaac	aaggtacgag	ctactcccat	acaacataca	aaaaagcaga	44040
gagagagtca	gagacagaga	cacacagaga	gagacagaga	gagagtctaa	agagagtcaa	44100
ggtctcagga	ctgaggggat	agtctactgt	agagcatttc	cccagcataa	acaagaccct	44160
ggattcaatc	tgaacacagg	aaaaaagggg	gggggggact	tcgattatct	caaattctcc	44220
ttttgtgac	acacccctaa	agtcactgcc	tacttccctc	accgccatga	agtaaagagc	44280
tgtttgcgct	tatgtctaca	cagtcctggc	tcccacttcc	tcctcccttc	tgcttctgtg	44340
ctcatctcct	ctgaaaccac	tgcagcaagt	gacttgtgtt	gactgccaca	cggaaactct	44400
cctcagtagc	aggcagcaga	gcagagctct	gtcttctcgg	agcttcttct	ctcttgtcgc	44460
cattttctcc	cacccttaag	taccctatct	tctctgtctc	tgcttgttga	tccttggacc	44520

cttttccttt	ctatgaacaa	aatatctcct	taaaggatct	cttctagttc	aggggtcccc	44580
cgccccact	gtggagaaaa	cccagggcct	tgcacatgct	cagcaggagc	tccatccagt	44640
ctctagctcc	atgacttaaa	gcattctctgt	gctgtcaaat	atacacttcc	agcccttacc	44700
aaaatattca	gtcaactcct	tgccattcaa	aatggatgac	ctcaaagcca	gagtcagcgg	44760
tgctatgact	cccagatcca	tccacttggt	agcccaggaa	tgaactcann	nnnnnnnnnn	44820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	44880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnngg	ttactgggat	ttgaactcag	gacctttgga	44940
agagcagcca	gtgttcttaa	ctgctgagcc	atctctccag	ccaccaccac	caccaccacc	45000
accaccacca	ccaccaccac	caccccacca	ccaccccgcc	ccacctgctc	attcctgatt	45060
ggttggttag	tttagtctgt	gagacaggag	ctgtcccttt	tctatagtgg	aaggngaata	45120
agaaactcct	gaaagtgaag	gcctacaaaa	cagccacact	tatttggtgg	aaaatactgt	45180
aaatgtgaca	tgtaaataca	tgctaaaata	attcggttaag	tcagtgaaca	accttaaac	45240
acagtctgta	gcctgaatta	cagacacgac	acgagccatg	acagaggctg	aaataagacg	45300
cctttgcaag	gagaagggca	gaagcttcca	tccttgctag	caaactctttg	ttccaagctt	45360
tatcagattt	tattgctttc	tctttctggt	tttctttatc	atatttggtt	atttggtggg	45420
ggaaagccta	atcttcatag	cccatgtatg	tgagcacttc	agcatatgtg	tgtaaacacc	45480
aacagcacac	gtgtatgagc	accacagcat	aggtgtgtga	gtaccacagc	acatgtgtgt	45540
gagtacgata	gcacgtgtat	agagttcaga	ggagaactga	gagagtccgt	cttttccctc	45600
tactgtgtag	gtctcagggg	tggaacttgg	gctcagccct	ggtggcaagc	tccttttatcc	45660
acagagtcac	cctgccagcc	cagctttctc	tttttctctc	tgttatgtct	atccactctg	45720
ttcaaggcta	actcactgac	tctgagttat	cagaactgct	tgtagagca	ggagtaactt	45780
tggaatctg	tgctggtagg	aacacccatcc	ccactcggct	tggtatgacg	aggggaaaaa	45840
aagcatcacc	aaggagtctc	accacctcaa	ccagcaata	tttacctcct	atacatggat	45900
aggtgggggtg	ggtgagcctt	gtgatttatc	gttaggatct	catgggagtg	attacagctg	45960
gtctactcca	tgacaaaaat	ggtgacgggtg	gctgacaaa	aagaaacagc	tacacctggc	46020
tctagttttc	tttctttctt	ttttcttttt	cttttaccct	acggtactaa	ggattgaacc	46080
caggaatgca	agagctctgc	caagttagct	acattcccag	atctgttttt	ccatttcttt	46140
ctttcctttt	agattttatt	tcgatttatt	tgtctatgtt	tgtgtgtatg	tgtgtatgct	46200
tatgtgtacg	tatcagtatg	ccatgggtat	acagaaacct	gagaaggcca	gaagagagtg	46260
tctgggttac	aggagtttcg	agctgtcctg	tggtttctca	aatgcagcag	caccaccacc	46320
aatcaccctcc	acccccaccc	agccttcgag	ttcaattctc	atcatcaca	aaacacacac	46380
acagacaagg	gacctgaaga	tggtcagca	ggtaacgaag	ctcgtgtcat	aagcctgaga	46440
acctgagttc	actgtctgga	acccgcgtaa	aggggggaag	gaagaatcaa	ctctatgatg	46500
ttgtcctctg	ctctcccat	gtgtgccatg	gaatgaacag	ccctcccaa	cacacatcaa	46560
gaataaataa	aactaaaatt	agcttagtaa	cttttatgtt	gaaagtgggt	tttacatgcg	46620
tgggcaacaa	taacaccgag	agtagaaagg	caagcatgta	tgtcactgaa	cagcattgaa	46680
gaaaaacaaa	acacatttcc	tgtacatcgt	tctggagtc	tgagttaggg	tttctatgct	46740
gggataaaaa	caccctgaca	aaaattaacc	tgggggaggaa	gctgtttatt	tcagctttta	46800
tgtctacaac	atgacctgtc	acccaggga	gtcagggcag	aaattcaaac	aggtcagaag	46860
cctggaggta	gaagctgata	cagaagccat	ggagctgctt	ctggcttgct	ctagcacaca	46920
ggagtactag	cctaggggct	gtactgccca	cagtgggcta	ggctctccca	cagcaatcat	46980
aaatttagaa	aatgcactac	aggtttgcac	acagggcaat	ctggtagggc	cattttctca	47040
attgaggttc	cttcttccaa	aaggacttta	gcttgacata	tggtgacata	aaaactagcc	47100
agcatattgg	gattatagat	attctcataa	aaaaaagaca	tttagattcc	cacataaac	47160
catattcaga	aattaactca	atgtgaacca	gaagctctga	aagtaagagt	taaaactatg	47220
aaaaattctt	acaaccatcc	ataacaaaaa	tctgatgccc	tcttctggag	tgtctgaaga	47280
cagctacagt	gtacacacat	ataaataaat	aaataaatat	ttaaaaaaat	atatgaaaaa	47340
tcaggctggt	gagatggctc	agtgggtaag	agcaccgcac	tgctctttcg	aaagtccaga	47400
gttcaaattcc	cagcaaccac	atggtggctc	agaaccatcc	gtaataagat	ctgactccct	47460
cttctggagt	gtctgaagac	agctacagtg	tacttacata	taataaataa	ataaatctta	47520
aaaaaaaaaa	aaactatgaa	gaactatgaa	ctacaagaag	tcaggaatag	ggctgggggt	47580
gtaacccaac	agaaaaacac	ttgcctggcc	tgcgtttggt	ctctagcacc	accaacgtag	47640
aaagagaaca	gcagaggatg	agggcatcct	gacttgagtc	aagtgaacaag	tgataatcct	47700
cgagacacca	aatcacaaat	gataaaagag	atcaacaagt	tggtgctttat	ctgaataaag	47760
agctgtgtcg	ttaaatacca	cgcaggaagt	gaagaggagc	tgagtctggt	aacacaggcc	47820
tgaaatccaa	gctactgggt	ggactgaggg	aggacaacag	ctagctcaag	gcccacctgg	47880
acgccagagt	taactcagag	agcagcttgg	gtagctttaa	tgagactctg	cccaggccag	47940
tgcaaggag	agatggctca	gtgggttaag	gcattactcc	tcttgcaaag	gacctgagtt	48000
caattcccag	caccacagtg	ggcacttaca	atcatccata	acttttagttt	caggggatcc	48060
aatgcccttt	tcacagtacc	aggcatgtac	acagtgcaat	tacatacata	catgcatgca	48120
tgcatacata	cacaggcaaa	acttacataa	aatactaagc	agataaatct	taaaagaagc	48180
cgggcgtggt	ggcgcagctg	tttaatccca	gcacttggga	ggcagaggca	ggtgtatttc	48240
tgagttcgag	gtcagcatgg	tctacagagt	gagttccagg	acagccagga	ctacacagag	48300
aaacctgtgc	ttgaagaaaa	taaaaaaaaa	aaagaaaaaa	atcttaaaaag	aaaaggagag	48360

gactggagag	atggctccac	agttaagaac	acttgttctg	aggtctacag	agtgagttcc	48420
aggacagcca	ggactataca	gagaaaccct	gtttcgaaaa	acccaaacca	aaacaacaac	48480
aacaacaaca	acaaaaccac	ttgttcttac	agaggacttt	ggtttgattc	tcagaatcca	48540
catgatgggt	cacaaccatc	agttgcaggg	atccaagggtc	ctgtcttctg	tgggcaccag	48600
gcataatgt	ggtgtacata	catgtataca	ctcatataca	taaaataaaa	agttttaaaa	48660
aggaggctgg	gtttgtagcg	cagaggtaga	ggtaaaaaga	ctctagcttg	tttaatgttg	48720
acatgaaaaa	aaaaagacat	ttagattcct	gcatacacacc	atatccaaaa	attaactcaa	48780
tgtgaatcat	aagctctgaa	agtaagaata	agcctagtat	gcactgtaag	gctctggggt	48840
cactccccag	cactgcaaaa	gatcatgaaa	ccagaaatgc	agatcctctg	aaccacagca	48900
tgggaatgta	actcagccga	tgcagtgtctc	acctgtcgta	tacagagcac	aggataaatt	48960
gatttgtggtg	gtgcataacct	ataagctcac	tacgtggaaa	gtagaggcag	gacgaccaa	49020
ggttcagtg	catccttggt	cacatagaga	atltgaggcc	agtctggtct	gctggtctat	49080
ttggaatgct	gtctcaataa	ataaaagaaa	gaaagaaaaa	gaaaagaaga	agtcctatga	49140
ttgtcttaac	ctctgacctc	tgtgttcatc	aagtctcctc	ctcaggaaact	cactggtcat	49200
cttgtgaaaa	cctacccag	agtctctgtt	cagaggaccc	aggctccagc	tgtggttacc	49260
acataggatt	tttatactag	aaaaataaaa	tgaataagta	tgtatttttt	aaaaagggtc	49320
agagctggat	atgggtggtg	ctagttatag	catccagaa	tgagacagga	tagccatgag	49380
gttgagaaca	gctagactat	acgggtctcaa	caaacaaaag	taagggatct	gagtagatga	49440
ggttttaatt	tttttctttg	tgtttgttac	ctaactgtga	tggttgtttt	gaatacatgc	49500
atgtctgtgt	atcacttggtg	tgcctgaaac	ccaaggaagc	cagaggaggg	catcgggtcc	49560
cccgggaagta	ttattacaga	aggttgtgag	cagccatgtg	ggtgctggga	atcaaactctg	49620
aaagagccac	ctcgggctgg	agagatggct	cagtggttaa	gagcactcaa	tggctgctct	49680
tccagagggt	tggagatcaa	atcccagcaa	ctacatgggtg	gctcacaacc	atatgtaatg	49740
ggatccgatg	ccctcttctg	gtgtgtctga	agacagctaa	agtgtactca	aataaataga	49800
tcaaaaaaga	aaaaagaaac	agccacctct	ccactctccc	tttttaaaat	cctcttgctt	49860
ctgtccctta	atgttaataa	cacagggtata	tgatactatg	ccttgtttat	gaatagaaaa	49920
tacacgtgct	aaagcaagtg	tgaaccttaa	atacattatg	ctgagtataa	ggagtgaagt	49980
gcacacaaga	cttttctgct	caagagtatc	tgtatgaagt	attgaacatg	tgaactctga	50040
aatcgggagc	tgaggaagat	atggggagtt	ctaattggcta	caacatttct	ttttggaatg	50100
atgaggatgt	tctagaactc	aaaaatggtg	ataactcagc	atatatacta	aaactcattg	50160
aattgtacac	tttaaatgaa	tgaataaaaa	cttgtctcag	taatgtggtt	tagaagatgt	50220
acagacatgt	gtgtgtgtgt	gttaaaacat	ttcttggcat	ggcaataaaa	atacagtttt	50280
agccaggtgg	ttgtggctca	aaaaataatg	ataataacaa	taataaaaaat	aatgaaaaca	50340
gaggctggag	agatggctca	gcgggttaaga	acactgactg	ctcttcacaga	ggctcctgag	50400
tcagttccca	gtaaccacat	ggtgggtcac	agacatctgt	aatgggatct	gatgccctct	50460
tctgatgtgt	gtctggaaac	agctacagt	aaagtcatgt	caaggacttt	acaatagtga	50520
ccatgataac	attgaageta	gacttgctac	tactgtctgag	tgtgtctgct	ggctctttct	50580
aaggagtaat	gttagctttt	tgtcctaaat	cttttctctt	cttttctctt	ctccctctgc	50640
tgtttttctt	tacccctctt	ttactttgct	ttccctctct	atctcctctc	ttaacagagt	50700
tgtcctatgc	agcccaaagt	ccatcttctt	gcctcagcct	ccccagtggt	gaaaaatact	50760
ctttccacag	gttatgttag	gagactggag	tctgtctcagt	cggggaggga	gcctgggtca	50820
agttctgagc	tcaattcctt	ttctttcttt	ctttctctct	ttctttctct	ctttctctct	50880
ttctttcttt	ctttctttct	ttctttcttt	ctttctttct	ttctttcttt	taagacaggg	50940
ttctttcttt	tacctgggt	gtcctggaac	tcactttgta	gctggcctgg	aattcagaaa	51000
tctgcctgcc	tctgcctccc	aagtgtctgg	attaaagggtg	tgcaccacca	ctgcccagcc	51060
ctgggctcaa	ttcttaacat	tgtggagaga	aaagtattgt	agctgttctg	gccacctgga	51120
attactttgt	ttctgatctt	ttgtgtcagt	caaactcttc	tcacccatct	ttcctcgtca	51180
ggctataata	tagactctcc	ttgcaatact	tggaaatgct	ctacagtcag	ctacatcctc	51240
agtcctgctc	ctataattttt	tcctaagctt	ccttctaagg	tctttattgg	tttatgattt	51300
acacagaaca	tttttttttc	ttgtctatag	catgcgttag	agtgatcgtt	gccagataga	51360
ggaaagagaa	atgagagaa	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51480
cagctactga	ttcctcctcc	tccctcctcc	ttcctcctcc	ctcccagcc	tcagtctctg	51540
ctcatcttgg	acttctgcgc	atgtcctcag	cccagacctt	ctgctcttgc	ttctcctctc	51600
cccagcagcc	cccagttctt	cttctctgaa	cttctgaggt	actctccatc	acctcctttg	51660
gctcctgctc	tgattgggtg	cacctgctgg	ataggcttgc	tcctgactcc	actgttctgt	51720
tctcaattag	ggacctcac	cctctgatat	accacacatt	tcctagtgtt	ctccacctcc	51780
cacccccacc	ctatacgcac	atacacactt	agctgcatac	ggatcctaca	ccagggaactt	51840
cttacccttc	taactcctcc	caccggacac	tgcggaggga	cactggggct	ccagagggct	51900
attgccacac	ggacacacag	gagatctcat	caaggagatg	tgcctacccc	agagggtagc	51960
tctcaccatt	cacaagcaca	ccacttctgc	ctccagcttc	tactctctcg	caggaagtág	52020
ccagcccggt	gccaagtatc	cccaactaca	tcccaaaaat	tctcagacac	tgccagcctc	52080
cagctgtcag	cctggccccg	gctggcgggc	gcctgtctct	ggcatagcga	ctagggtgta	52140
attagaaacc	cgctagctcc	ctaattgcca	gttctgagct	gtccttggtta	ccggctgccc	52200

gaggcacaca	tagaggaaaa	ggctgagagc	tgagccaggc	tggcatggag	gtagccctag	52260
tagacctaga	gaggactggc	atgtggccag	ggaccaaacg	tggcacagag	agggctcagt	52320
gcaatctgcc	ccgtgggtgc	ctcccagcca	catccatttg	cccagaactg	tgacgtcaaa	52380
ccagcccggc	ccattcattc	tttattcagg	tggcataaaa	atcactacaa	aaactttaca	52440
aaagagtctt	gggagctaaa	gggtcccttc	cttgccctcag	tccccaagat	tccctggcagg	52500
ggaggacaag	agagagaaga	aggaggaaga	ctcctggcag	tgttggcatc	tccaaatacc	52560
agaggggtga	cttgggtgac	aggacacagg	ttggggacct	gaatgtcttc	agcaagggac	52620
actctttag	ggtaggtcag	cctccaacca	tgaagtataa	caccaaggcc	agtctaagct	52680
tgggagacca	acacttgtct	ctccttttcc	cacccagggt	gtctggaata	tgtctaaaga	52740
tggcctctcc	agcctctgct	tacaaatgtg	gagggaccct	aagttaggga	cttgccctaac	52800
ctacctctag	ccaaaactgt	gtccacaagt	gccagccac	aaaagatcac	cccctgagcc	52860
ccttgggaag	aaatgaagat	tcccattgct	tgcccttctc	caggcccccac	cccacatgct	52920
gcaagagaac	agcttctaca	ctgggtgatgg	tccttccggg	cccacccat	cccacaaagc	52980
tgggttagaaa	gagtcacagg	agctgagagg	ctgatccagg	tggggactca	ggatgctgct	53040
gcccagggcc	cctcctcact	tgggggagct	gaactggggg	tagtcttctc	ccatgcgggg	53100
tgcaagtttc	aagtcaggac	caaaggctct	gcctccatgg	aagtcagctt	tgtcattctg	53160
gcctatgagc	ctggtgtcag	gggaatctcg	ctgttctctg	agctggggca	ggcgctggg	53220
gttaggggttc	ctcacactgc	ccacaaagag	gggcacgcct	atgggtgtct	ccatgatgaa	53280
gaagaggaag	ggtcgggttca	cagtgaagga	ggagagggac	attcgattca	tggctacgct	53340
ggtagctgcg	gctgcctcca	caccagcctc	gctgagctcc	atggtagact	gatgttgac	53400
gctagacacc	accagattct	gctcagagat	cccacgaagg	tctgggccct	ggaacaattc	53460
ctgcaggcct	gcccagaaca	gcagatgact	ggtcagtgct	gccccaaagg	tatgtggatc	53520
tgtctagcat	ctgggctaaa	gggaacactt	gaacccagcg	gttgattgga	atctgttaga	53580
cctcagctca	gacaacactt	ctagaaacct	tttttttttt	tttttttttt	ttttaaatca	53640
ggatctgogc	taggtacagg	acagaaagtc	tagaggagca	tatcaaatgc	tcccatccag	53700
gaagcagggc	cacctctggc	tcaggcacac	tggcagctcc	cgtactctgc	ccagaccacc	53760
taggggcacc	ctatccccaa	gctccttacc	cagttggctg	aggggtggca	ccaggtccag	53820
ctgctgttgc	agatggagtt	taggcagcca	caccttgggt	ggcctctcct	gcagcgaggg	53880
atgggtacaga	gtatcccagg	tcagggttgg	tagtacctcg	gacacgttcc	actcaaaata	53940
agtgggcac	acgaccacaa	agctcatgtt	gttctttaaag	gggaaatgag	ccacctacag	54000
ataagaaaag	gagagaaat	gaggaccaga	cagcacctgg	acctgtctgg	agtctgggcc	54060
aaaattactt	ctgtactttt	gagacaagag	ccagaaattc	agggtagtag	tgctttcact	54120
taactggtga	agtggaaata	taccacttac	ccctttgcaa	ggtgacatgg	gaccaaatga	54180
gataatgctt	ttacacctct	ctgtgtgcac	acataagcat	atatgtttgt	atcgggtgtga	54240
gtgtgtttgc	tcatgggtat	atggagtcag	aagtaggtaa	acatcagtcg	tcttcctaca	54300
ttgtctctca	cttttttttt	tttttttttg	gtgttgccat	ctttttgttg	ttgttatttc	54360
aagacaggct	ttctctgtgt	agccctggct	gtcctggaac	tcactctgta	aatcaggctg	54420
gcctcgaaat	tgcagagacc	cacctgcctc	tgccctctga	tgctgtggat	ctaagatgtg	54480
tgtaaactaca	catagctccc	tcttttttgg	acacagggtc	tcattggatcc	caagctggct	54540
ttgaaatgac	tggttggggc	tggagagatg	gctcagcggc	taagaacact	gactgctctt	54600
ccaaaggctc	tgagttcaaa	tcccagcaac	cacatgggtg	ctcacaacca	tccgtaacaa	54660
gatctgactc	cctcttctgg	agtgtctgaa	gacagctaga	gtgtacttac	atgtaataaa	54720
taaattaatc	tttttttaaaa	agagaaagaa	atgatggcta	catacttctc	tctcgtctct	54780
ctgcccgaag	tgtcgggatt	acagagctgt	acaacaagcc	caagtttgtt	gtgttttaga	54840
catgctaag	tatcccaggc	tgtcctcaga	ctctctatgt	aattcagaac	gaccttgaac	54900
ttcttttaag	gtttattttt	atcttatgtg	tatgggtatt	ttgcctgagc	atttgtctgt	54960
gtaccgtgtc	cttgagctac	cctcacagtc	cagaggaggg	caccatttcc	ccctgaactg	55020
gttgtgagct	gcatgggtgg	tgtcgggaat	caaaccctgg	tcctctgcaa	gagaagccag	55080
taagtactct	taactgctga	gccacttctc	caccttgagc	ttttcttctc	cctatctcga	55140
tctaaaagta	ctagggtatg	cggatgtgcg	ttcatgtgcc	tgggttatgt	gttgctaagg	55200
gttgaacaaa	gggctttgtg	catgccaggc	aagcactcaa	caactgagct	acacatcccg	55260
acagactttg	actcttctag	tagtagtgtc	tccactacag	cctgagttct	ctatctgctg	55320
tcagcaagct	gtacaaacaa	gctatgggcc	ttcctgtcct	tgccctctcag	tctctccgc	55380
aggtggggct	actggctttc	aaaatgacct	atagaggagc	cacagcaaac	agtaggaagc	55440
ttgcccctcg	tctttcaccc	tctcccagag	agtcagctat	aattcgagtt	tttttttctc	55500
ctctctctct	ttaaacagga	tctgggtatg	tggccctaac	tatcttcaac	ttcagttctc	55560
ctgcttcaac	cttctgagtg	ctgggattat	ggtgtaagcc	accacactca	gtcacacaaa	55620
cctttttttt	tttttttttt	tttaaagaat	ccatgcagtt	aggacagcat	ggaaatgacc	55680
aggctcaggc	ctccctgggt	accagcataa	tgccctgagg	cgggtcctct	gccagtgagg	55740
ggatggaaag	atggagccag	aggatcttcc	ctctctgaac	ctcaatgtcc	cacagtgaga	55800
cactcatgtc	cactgggaga	tactgtagta	ttcaagggaag	aagcaacagg	aaggtgagag	55860
ctaagtggag	ctgagcaggc	tcgtatcctc	tcaccacggg	ctacagagaa	gtctggctgc	55920
cccctccaca	tggtcctctc	ctgcagaact	ggcaatgctg	ggcccggctt	gcccagtgaa	55980
actaaccaac	agaatggatg	agcatgtgtg	gtgccacaca	cctgggaccc	cagcactcag	56040

acagctgggg	cagaagggtc	atgagtccaa	agcgaacttg	tgtaacattg	tcagaccctc	56100
gaacaaacaa	aactagcccg	tcctgtttatc	tcagccacag	atgatggggc	caaggatcag	56160
tactctagcc	aaggagtcac	ggttaggcta	gaagcaagg	aagccttagc	tgagacagct	56220
tggcacggag	cttcatccaa	tcagaatggt	cagagcaata	agctttgaaa	cccgaacttc	56280
atctatgaag	cactgtgtgg	gaactcctct	cttcccttac	gagcaggggc	ctggtcctct	56340
tgggctccgc	taaaacccca	gcacagagaa	cagttacctg	gcacgtgaca	aaaactcaat	56400
atattttctt	tgaggagatg	aacctcaaag	aagctgtgtc	ctggatagac	acagcataat	56460
aaacccttca	ggagctacct	accaggggac	cagactttac	ctcccagtag	caggcctcgt	56520
ttgccagcca	aaggcaaagt	ccagactgac	ctgtatctca	ggttgctcca	gcaggaacca	56580
tcgaagagga	tatgacaccg	cgtgcatcat	gtccaccgac	actgtgaacc	gctcatccag	56640
gtggaagaaa	tctttctggg	tgaggctcgg	gtcaaaacttg	gtcctccaga	aacctgcagc	56700
caggcagagg	gcaggagcca	tgtaacataa	aatcagcctn	ctgcctgtct	tgcctagaac	56760
ctatnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56880
ggaaaaagga	atcttaaatt	agaagatgcc	ttgataagat	tggtcatgtg	gtatgtctca	56940
ctaattgattg	atgtggaaag	tcacgaggga	tggtgtcacc	ctgggcagat	ggcctggggg	57000
atataaaaac	acaggctgaa	caaaccacaa	agcagtagtc	ctcaatggct	tctgctttag	57060
tttctgtctc	aggttcctac	cttgacttcc	ctcagtgaa	gcatgtcaca	tgagagttgt	57120
aagaggaaat	aaaccctttc	ctcccccacat	agtttttggg	tatgatgtta	tatgtcaaca	57180
acagaaacta	taactaatat	agttgggttt	ctttttttgt	ttgttttggg	ttgtttttag	57240
acagggtttc	tctgtatggc	cctggctgtc	ctggaactca	ctttgtagac	caggctggcc	57300
tcgaactcag	aaatccacct	gcctctgcct	ctgctcccca	agtggtggga	ttaaaggcat	57360
gcgccaccat	tgccctggctg	gttttctttt	ttttttaata	catttataat	gcatttttaga	57420
tttaaaaaaa	aaaatggcca	tgccatataa	tataaaaaga	agtgtttaca	aatcaccatg	57480
tgcccttgcc	ataaattatg	taaaaatttc	catatggaca	tcagtctcaa	gcttacaatc	57540
tcagcactca	tgagcctgag	gcagaggcag	gaggatggg	agctcaaggc	cagcttagtc	57600
tacataacaa	gatcctgtcc	aaataataac	aacagtaata	atttcataca	tagaactaga	57660
aggggcccact	gcaaagacag	tatgacaaaa	ccactggccc	tgccctaattg	tattttaaat	57720
aactgtcctc	ctctctgtaa	ttttcagttt	ctaattttta	cataactacc	atgtattctt	57780
tttgtaattt	taattagttt	tttaataata	gaaacaagct	aagtgtctaa	aatattttca	57840
tatgaacatt	ttcaaggcac	ttgatacata	cctcagattt	gccctccagg	tgagcagtag	57900
caattacgtg	ccaccagcaa	tgttagcttc	cttttttccc	taccatctga	ttctgtttca	57960
gtctattcgt	agttctgac	ttgttatatc	cttttttatt	gtttccctgg	gttccaacac	58020
ctcccagttg	agtgtttctc	ttgaatttca	ttagcagctg	tttcattaat	ggcacagaag	58080
aaggattaca	gtgttaacta	ggatagactt	tgacaaaaga	ctatgagAAC	atatcttatt	58140
atctttgcat	aaattctttt	taatcaaagt	tcctcaaaag	cctctctctg	ttcccatctc	58200
agggagtagg	tctggccact	gatgagtgtc	caggccacag	tacagggtg	cgtgggtctg	58260
tcctctgtgg	aagggcacat	ctgtgttgta	acaggattcc	tgtcttaaca	agccttgctc	58320
aggctctaag	tggtcctgag	ctagctaact	gcccttggtc	ttcccttgat	taccagataa	58380
ctattcactc	ttctcatttt	gcagagcact	taccaggtag	ctatgtcctg	gaagtagcaa	58440
tgagtccttc	tattgttttt	cttttaactta	aatcccat	gaaatgcgcc	agggacactt	58500
caatccaagg	tacacttttg	ctaaagaatc	actcattttt	atatgcaaaa	tgtaacctat	58560
taactgcagc	tgatatggta	catacatatt	ctctcttcc	attatccact	aataggtagc	58620
taatgcgaaa	tattgagtaa	tttttaaaaa	tcaataactca	attttttaga	aataattaga	58680
gagacattca	actctgacac	cagcacccta	ctcagttcct	gagccttcct	ctgccggagg	58740
agaatctata	aataactcac	gaagctgaca	ttactcactg	tggtgcagtc	atttttttct	58800
gagaaaattt	tagcaactgt	tctaataag	cctgccagtt	atcagtagtt	gagaatgcaa	58860
gtcaactttt	aattatgcag	acgctgatta	ttcagacgac	aaattgttgg	tgccctgcag	58920
gctccttcc	gctgcctacc	tttaaccgtt	ctcagtgctc	attagcacat	gttccagaag	58980
gtaggetttg	gaggggagg	caggcactca	aaccagctaa	gcacttagag	aagctctgat	59040
gaaagatgtt	aatgcagttt	gtagaattat	tgactaaaat	tgagtcattt	ggattccctg	59100
tgaattgtat	ttacatgccc	tgccctgtc	ccccatagca	acagataata	ggattgtctg	59160
cagagagaca	acatagttct	tatatttaat	tttttctttt	gtcgaacatt	ttcacatgat	59220
ggttcgtggg	gtttcctttg	ttcattacat	ttgtatccag	actagttact	tctgataagc	59280
ggttagttag	gattcctggc	acgcggacag	tgacaccaca	gttgtctgat	cgtttccac	59340
ttttttacaa	aaccgtttgc	ctttaagagt	cagtgttttg	cacatttcac	ccagattatt	59400
ggaaatatta	tttccctcct	gcttaaaccg	aagctgtgat	cataatttaa	gcctttctag	59460
gtagccgac	ttacatgtat	catacctatt	tctggcatal	gtttgtctat	tacaaagacc	59520
tcgtaggtat	gcagttagaa	gcctctagtt	aaatgaaatg	ttgcgtgtgt	gatgaacctg	59580
gagtgggat	ggccttttgt	gtgccccaa	gctgtgtgt	ttcacacagt	tgttttctgc	59640
ctcctctggt	ctatcactat	cctgccactg	ccagaaaacc	ctgctgtgtg	ttcccccgct	59700
ggaggatctc	tgcttctgaa	cttctttggc	ctgagaaact	ccataaccaa	atcagttagc	59760
attttgttta	aagagcaggt	aggctgttag	agcttgggtc	ttacatgtct	cccaggtcca	59820
cttgccagcg	ccttgaccac	tgtaaacctt	tgtaaaccaa	ctcatctttt	gctgcctgtt	59880

ttttgggggg	tttttttggg	tttgtttaag	ccaagatcag	ttatatggcc	caggctgagc	59940
ctctcttccc	agcctctcaa	atgttagaat	tacaagcatg	catccctcag	catacctttc	60000
ctttgctttt	tttaaaatag	agttttgcc	tagcaacaga	aatctaacct	aactaagcat	60060
agccgtgcac	atgggtatgag	gaactcacat	atgtgtgaat	ggaagttcat	agagaccggc	60120
atcactgcct	agaggccctt	ttcttccttc	cttgcaagttg	tcgtgctagc	tgactgtact	60180
acaaaagagg	ttgtctgagg	cataagacta	cottcaataa	aacatgcaca	gacagtttgc	60240
ttctctgaga	tttcagagca	gtgactacct	tcaataaaac	atggacagac	ggtttgctta	60300
cctgagactg	cagagcagtt	tccaaaaatt	ttagacaaag	ggtaggatga	agaaggctgc	60360
ggggtttttg	acacacttaa	ggtgcgtaag	taaataaaact	gagctacact	gacaggatgc	60420
tcgttctagt	agccaaccaa	agagcagttg	aaccaaagca	cctagacttc	aaacatcggt	60480
gggagataat	cttaggagtg	ctatgcttct	gcgtccataca	agtattatga	aactgtctag	60540
aaagcaccoc	actggtaatc	cctttttgat	tatttttttt	ataaattcta	gtcttggggg	60600
tttgagtggc	acacagacat	aatgggttagg	cttcgggtgtg	tgctcattca	ctttgcttcc	60660
tggggaccag	agttttgcgat	gagtcagttt	ccatctgatt	tctgtcggat	ccggctgcag	60720
agccatgact	cagatgggct	tcaggcccag	ctgctcagtt	catcttctgg	ggaatagatg	60780
acaaggacgg	gacaaatgtc	ctgacgcaca	tttccttctg	ttccttgcact	tccagggtct	60840
aacgagagca	tcattaccac	cagcaggcag	atacgccttg	ccacaggcat	cttcctgtgt	60900
gtcagcctcc	tgaaccactc	ctgcaggccc	aacaccagtg	tgtccttcac	tggcactgtc	60960
gccaccgtcc	gggcagcaca	gaggatcgca	aaaggacagg	agattctgca	ctgctatggt	61020
gagccagcct	ttctttccac	tacctgtctg	tgcctcacac	ctcacatgaa	aaggataagg	61080
ggacaggaat	cagcagatat	gggcccagtg	cctctactca	tcctctgagt	ctttcctgga	61140
aagggaatag	catccttggg	ccaataaaaa	aggtcttctg	gctgtaataa	aaaagcccgt	61200
tgagggcagt	gagccatata	cctccatgcc	ttgtagacag	cctatcctga	aaatgagcga	61260
ggagcacttt	cttggtctct	ttcttctctg	cccagcagct	tggaaacgta	tccactttca	61320
cccgtgtttt	gttggttttt	ctgagatgat	agggcagagt	acccaacctc	atataggcta	61380
ggctagtgtc	tatcactgag	ccaggacccc	aaccagcac	caccatgcca	gtcacgtgat	61440
gactaggcca	gcccctcggt	agagtaggca	ttgactctct	tggtgtgact	aggaactgtg	61500
ggtaatctct	ctccagggcc	tcacgagagc	cggatgggctg	ttgctgagag	gcagcagagg	61560
ctgagttctc	agtacttctt	tgactgccgc	tgtggggcct	gtcacgctga	gacactgaga	61620
gcagctgcag	ctccagatg	ggaagccttc	tgttgtaaga	cttgacagagc	gctcatgcag	61680
gtaaatctct	gctgttccca	ggggcagggc	tccagctaaa	ggttgtcagt	cgccaggaga	61740
accattcctg	cttccttctt	tgtaactcct	cctcatctgt	cgcccgggtcc	gtcagaaaac	61800
acaggtgtga	tttccctaata	tttcccttat	aagtgcacac	aaatcttaaa	ttacacaaag	61860
ggacaaaaaa	aaaaaaaaaa	aaaaaagccc	tagaaattta	cttgctcaaa	taagtcatca	61920
aaagtgtgtc	atcaggccta	gcacttgggt	actggtaacc	ctagcactca	ggaggctgag	61980
gaagaaggat	ctcaagtcgg	aggccagttc	caagtgcac	cccatctaag	agatcaccat	62040
tccaaggagc	tatttcagag	atggtttaat	ctggggaccc	agattgtgga	ttttctgtct	62100
gttcaatttc	atctctctgt	gctggcctca	tcagacacac	tctgtagtaa	ctgtgggaaa	62160
atccgaccca	catagttttc	cctcagcctt	tgacccagag	ggaagagcca	cagtggagag	62220
catgagagca	gacccttggg	tgctactgcc	aggtaatggt	gtagacactg	gagtcttcaa	62280
cattcatgcc	ccaatgcaaa	atggtctcca	caccagagca	tggcattctc	attagaaata	62340
agtaaatgga	attggctgtg	ttgaaaaattg	ttaagccaag	ggtaagaat	gaagccttcc	62400
ccagcatgtt	ttgttttgtt	ttgtgtttta	ggcagcgtct	ctctgtgtag	ccttggctcc	62460
tgccctctgc	tacctctccc	agggtgtcca	ccatgctggg	cctaagcgcc	ctgtgcatta	62520
gtgtccctc	gatectgtct	actcttgaga	cagtcttctt	tctactctgt	atccccagat	62580
aacctagagt	tcacttcaga	gcccaggctg	gcctcaaaact	tgagatcctc	gtgtcccagc	62640
ttctcaaatg	cagtgatatt	tacaggccta	cacctggctt	tccttgatag	attcctagta	62700
agatgattat	cctttgagcc	atatctctct	tctgcttctt	cctctcttcc	tgcagggttg	62760
atctagaatt	tattctaaag	ctgactggcc	tcagaattgc	catccttctg	cctttagnnn	62820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	63000
tgcccttaaga	gttctttgtt	tatgggaatt	tatgggaattg	gtgtgtgtgt	gtgtgtgtgt	63060
ggagtctctg	ctgttggttt	ttaatcttta	ttcatcttgc	tcagtctcct	ctttccaccc	63120
gtgtgtatgt	gcgcgtgtct	aacttgcaaa	attgcaaaat	ggtgctttac	tgagccatct	63180
tgtaggtect	ggggatcaga	ctctgttagg	cttggtggta	ttctggcttg	ggtgtgatag	63240
tacaggcccc	ccatggacaa	ctttttcttt	gaaaacctgt	aggaggacta	acagaattgg	63300
ctcacacctg	tgacctacc	accactcatg	gctaactata	tactaccac	atggcaagac	63360
aagccagcct	ggactacaca	gtgagtaaa	aaaagtcaat	taatttcaca	taaagtcaat	63420
ccggttttaa	aacactgggc	aaggtgaaac	actgtatgca	ggtagtagt	tggtttcaat	63480
agcttcatta	acggcctagt	tatcttttaa	gtacttttagg	ggctggagag	tcagttaaga	63540
tttattactt	tttctctgga	acatttaaaa	actcccaagc	acccagggtg	caatcacac	63600
acagtggctg	ctgccaaagg	actggagttc	accctcacag	actcacagcg	agtggaaac	63660
tgtctgtcat	ctaattctag	gggatctgac	gaaataaaat	accaggcaag	gtggcacacg	63720
caatgtacat	aaaataataa	ttaaaaaaat				

cctttaaccc	cagcactcag	gaggcagagg	caggcagatt	tctgaattcg	aaggcagcct	63780
ggtctacaga	gtgagttcca	ggacagccag	ggctatacag	agaaaccctg	tctcaaaaaa	63840
aaaaaaaaaa	aaggacttta	aattgggctg	gagagatgga	ttaaaagcat	tggctgctct	63900
tcccagaggt	cctgggttca	attcccagca	ctcaaattgg	ggctcacaac	tgtctatatc	63960
aacgcaatct	aacaccctct	tcaggcatgc	aggttaacatg	tagacaaaac	atccatatgc	64020
ataaaaataca	taagtaaatg	agtcttttaa	tgtatactag	aagctgggtg	gtgggtgcatg	64080
cctttaatcc	cagcacttgg	gaggcagagg	cagggtggatc	tctgagttgg	aggccagcct	64140
ggtctgagta	aatagagcct	tgtactttcta	cttatcacta	cagttacatt	ttataacttt	64200
gggcccctagt	gcttccattt	tccactgttt	gcttaaccac	tggggcctga	agcttttgtg	64260
ctgacacttt	tgttcgctaa	tcacactagg	accaatgggt	tctacactcc	atcaccatca	64320
acacaaacaa	aacaaaacac	aacactacgg	atcctggcat	ggtggaacat	ctttagcccc	64380
agtagctggg	cttgagttca	aggccggcct	ggtctacata	gcaagttcta	ggatagtagg	64440
gtagtctttt	aaaacaaaac	actattttat	ttatgaacaa	aacatgtaaa	gaaagaaaaa	64500
aaactgcaaa	tttatctatg	aatgaagtct	aagtaatact	tcaatatattg	aaatagcttt	64560
ctaaaatatt	tttattttaa	gaaaactcag	caaattattc	aaacaacctt	ataaacgttc	64620
gttataaaaag	taaagaatta	tttgcaattg	ccttaacggg	ccaaggtggc	agcctcttaa	64680
aattcagaac	aatccaagct	tcacattcca	gttcaacatt	tctacagccc	taacgtattc	64740
aaatacctcc	attctgacaa	ctgtttcccc	tcttcttttc	ttctaagctg	cttagatgtc	64800
tgtcccaggc	ttttcatgat	tttagtcatt	cacacaacta	gcaaacatta	tctagggact	64860
aaaacttgcc	agatactggg	atatcaccct	aaagggggac	tgaaagtagc	tgcaggctac	64920
agtctctaca	atctcctgaa	tgaatacaca	agtagctaa	atttaccaaa	taaacatgta	64980
cacctgtgat	gattgctagc	tgtactagca	gaagctaaac	actaaatcta	gaaactcagt	65040
cctccaacta	gccccttget	cggcttcagc	ctcattttta	caaaacaggg	aaagagtttg	65100
gaatgttgcc	caaagccata	cataagtga	caaaaaggag	ttggagtctc	caaatgcatg	65160
gatttgggct	agttactttg	ccaaccaact	cagtaacaac	tgagctgaac	aggaacactg	65220
tggtagcaaa	agaaactgga	actatcaatg	gcctctagag	caaaaatata	tttaaaaaga	65280
aaaaaacaaa	caaggcctgg	caaggagact	gtgagaagag	tgtgctgact	gaaattgact	65340
agttcagcca	acaaaagact	attccagggc	tggtgagatg	gctcagtggg	taagagcacc	65400
cgactgctct	tccgaaggct	aggagttcaa	atcccagcaa	ccacatggtg	gctcacaacc	65460
atccgtaaca	agatctgact	ccctcttctg	gagtgtatga	agacagctac	agtgtactta	65520
catataatca	ataaataaat	ctttaaaaaa	aaaaaagact	attccagtgg	ggatggaaaa	65580
gttaagtgtg	gagttaaaaa	atacttcaac	tggtgatgga	ctaggtgtcc	agagtccggc	65640
aaaaggatgc	tctgtggtag	aggtgcctgc	tgtgtaagcc	cagctacctg	agctcaatcc	65700
acagaatcca	cagcggagtg	ggaagagaaa	caacgtccca	gagttgtcct	ctggcatccg	65760
acgcacattc	gccatcccca	agatgtcata	catatgtgta	catactacac	actggcgcac	65820
gcgcacacac	actctttttt	aaaatttcaga	cttagaggga	cataaaggat	ttgctctgat	65880
atatgttcaa	ttgaaaatga	ctttgaagat	agagggcaga	tcgaagggaag	ctcagcagga	65940
aagaattaat	aacatgcagg	tgaagggcta	taaaactagt	tgcaaggggc	cttggctcga	66000
caaaaaaatc	tatgggggtt	gccggtaaaa	taaggaaaaa	ggtgtcaaca	tgaaacacag	66060
aacactagca	agagaggagt	gttagcagaa	agaagccaac	aagctcaaac	aattaggtcg	66120
gctgaaaaat	tttaaaatgt	cttctgattt	ggctactggg	aagccactgg	tgacttcggt	66180
cagcgttttc	tctctcgtga	ccagagagat	gtctagtagc	aataatgagt	taggagtagt	66240
taaaagaagt	aaaacagccg	aaaacaagtc	caaaaagttt	ggggtgatgg	agaaagggag	66300
gaaaacagag	cgcgcgaaga	tagacagcgg	catgtttatt	tgtcttgttt	tcttagatgt	66360
aaaacaaact	aaaaaactcg	tgagttcttc	tgccagtacc	gggttgccct	cagcatcctc	66420
tgatggctct	agagaccccg	ggatgctccc	ccgcggccgt	ataatttctc	ccctgacgct	66480
ctcccgatcg	acagcggtc	cctccccggg	tcctctttgc	accgctccaa	ggcgcgctg	66540
ctagggccat	cgagcccgtc	cagggtcgtc	tccttacctc	gatggccccc	tcgctcaggt	66600
gtcccacat	ggctgcaccg	ctaactcccg	cgtcgcgct	cttgaccgc	ctgagcttct	66660
gtgccggggt	cccgcgggct	gtccaacgat	tggttagagc	aactgtgcgt	gccgatccgc	66720
ccccagcgtg	agcgcgggtg	gaggggagg	cctagacgcc	gatagccacc	gcattggcta	66780
ccgcgcggca	ggcagagcac	gtgactcttc	cgaggccggg	ttcgaggcct	agtggcggga	66840
tggcgggacg	tgagggcggg	gcgctgggtc	gcagtgcgcc	tgtgtcagcg	cggtgctact	66900
gagttgttcc	cccgccagct	gtcggaactt	tgcccgccca	gtcctttggc	ggacagacag	66960
aatggcaacc	cagggaaacag	tcggagctct	cccctggtaa	ctgctgctaa	atatagtcag	67020
agcagtgacc	tgggtacttc	ttcacgcagt	gcgtgcccg	cgccgtgcc	aggccacag	67080
cttggcactg	tgggataaac	aaggtaaatc	agactcagtc	tccgccctct	tgagttccac	67140
ctgagagttg	tggccgcaag	gaacccagcc	tcaagtagtg	tagacgcgat	atgggccaca	67200
catgtggagc	tccagagtgg	gggtcaaaaa	tcaatcaggc	tttcgagagg	cgatgcggtt	67260
tgaactgagt	taaagtgtgt	gtagaaatct	gtcaggtgga	ttccagttag	gtagtgatg	67320
ttcctaaaag	cccaaatggc	ctatgcaaaa	gtattggaga	gcctggcggt	ctggctggct	67380
ctgatctgtt	tgtaatccca	gcctttggga	tgtagaagca	gcaaaaagtc	aaggtcacc	67440
ttaacaccgt	tgagttcgag	gtcaacctga	actaaatgag	accctgaaaa	atcaaaattt	67500
gggacccagg	cgtggtggca	ttcgaggtaa	aagcaggcag	atctctgagt	tcgaagccag	67560

ccaggctaac	ataagatccg	gtctcaaaaa	aaaaaagtaa	taaaaataaa	aagggagaga	67620
ggctatatga	actgaaagaa	agacctggag	atcaaaacag	aaaactgagc	cgtctaagaa	67680
atgaaaatat	ttaaacttcat	agttgctgga	gtaagaagtc	tggaaaactt	tgggcaacta	67740
aggtaaacag	gtctagaaag	actggaatag	tagccatcta	ctggatattt	gatctctggt	67800
tgtacaacca	caacctacta	tagtttctca	aacagttcca	aagaatatgt	ctgggtgaat	67860
tggtaccaca	ccacagatta	actctccttc	agcatatcaa	cagctataga	aaaccccaga	67920
agaaatgatt	ttggttgcg	gtcacttggt	aggatgaaat	ctcgattttc	tagaactatg	67980
cattaataga	aagctgaatc	ttcatgttct	gactttacag	agctgcggca	gcatggatct	68040
accggtggat	gaatggaggt	cctacctact	taagaagtgg	gcttcactcc	cgaagtctgt	68100
gcaggacaca	atttctacag	cagagacttt	gagcgacatc	ttccttccct	cttcttccct	68160
tcttcagtaa	gtgaatggaa	acttcaggga	aatttttggtc	tggaaaatgt	tctgccttgt	68220
catttggtct	gaatatctct	tttttatagg	agagagtgc	tttatattct	ttatagtatg	68280
gggcatttag	cagttactgt	tggttttcac	gtttctccct	agtctgtgat	tactagaatg	68340
ggtaggcact	aactgctttc	ctcttttggtc	atgtgttata	cttaaggaat	gtagtatctt	68400
gctgtcgtcc	cagtgtgtgc	actcatagga	tctggtgcag	gttggtgtagc	tgcccctaga	68460
agctcattca	gtcctaattg	ggagaaagaa	ccctggcact	tggttagttg	agaccanana	68520
cttctcaagt	tctnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	68580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnntttagtt	68640
tccaagtgcc	actttactgc	aatgtgtcac	cacatccaga	gttctgtgtt	tgtttatttg	68700
tctgtttttg	agacagggtt	tctctgtgta	gccctggctg	tcctggaaact	cactctgtag	68760
accaggctac	cctcaaattc	actgagatct	gcctgcctct	gcctccagag	tgctgtgta	68820
cactaccacc	acccatccag	ctttctatat	tggttttctc	atggcgtttt	aaaatagcca	68880
ttatacgtgt	gtttatcatc	taaaagtcctg	gtcccaaaaag	gagatgagag	gggctgctaa	68940
ggtgaaaagg	attacaaaag	cttatcaatt	ctgttcaaaa	attaaacctc	agagtgggat	69000
tagctgcttc	tttcattaga	attgctatca	gaattcactc	aggccttggt	tgctgtgtgt	69060
attgaagaag	tctcttcttc	atcagggtcag	tgactcctta	gctcaagtac	atgcaatatg	69120
cagtattgat	aactgttctc	gcttaggaat	aaaaatagaa	ctgacttcca	cagggaaatg	69180
atgtgctgag	ctgtagcaac	gaatcttgca	caaactctgt	cagcagggac	cagctagtct	69240
ctcgctgca	ggaccttcag	caacaggtct	gcatggccca	gaagcttctc	cgaaccggta	69300
aaccagtgga	gattggctcc	ctcgccctag	gcctcagccc	ttcccttggt	tattttggta	69360
tcaccttgcc	ttactgagca	gtcctcaata	aatgactgag	gacttgaatt	taattatccc	69420
agcaccagcc	acaagatggc	tatgtaggcc	agttagacca	gactgtgacc	agctgttact	69480
ctgggtgccct	tgaaagtctt	cctgatgggt	taagctgtgt	ctgctgcgcc	agatagttct	69540
agcagctcga	gcaccagaaa	ggctgtctga	cttccatggg	ctttgtgtgg	ctccagaggt	69600
ccaatgccat	catctgattc	ccagcttaag	gacctaagct	ccgagaaggt	tgctctgccc	69660
tcagcagcag	cagcaagtcc	tgagtgtctc	ctgggctcgt	ggtgtgactc	aggagttagag	69720
ctcggtagct	agcctgagct	gagagctgag	agaaagaaag	gactcctctc	ttttcagaaa	69780
gggatttgca	gaactcgatg	ttagacctg	acatggtagg	aatctgtttt	gactattcta	69840
gcctagattc	tgaagttgac	ctttagccta	gagtcagaaa	aactaatgat	tacaggagga	69900
atgtagagtt	ggttggttaa	tggttggttg	aaaatggatg	ttagaagccc	agggtaaaatg	69960
tgaggaagcc	tcatctaaca	cctcttttac	tgaaagagaa	aacataagca	accaacagct	70020
tccctggaat	gcccggctgt	tgactccgtg	agataaagag	gcattttcac	tttgacctaa	70080
ccgatagaga	ccttgcaacg	tggtctctcg	tgccaggac	tagatctgta	tctgttgtga	70140
ggcatttttc	ctttgaaatc	ccattcagca	ccgtactggg	ggtgttgagg	tgccgggagg	70200
ctgctgagag	cttcttgtca	gcagagcaca	ccttaggtata	ggaaattgaa	gatggcctgg	70260
cccaggccca	tgctacctta	ggtatgctac	tggttggttc	gccggagttc	tccttccctg	70320
ccgtgtgttc	agtgcggccc	ttgccttgct	gccccagag	tctcttgcca	tctgaattga	70380
cgctcttctc	cctcccatte	tgcatctctc	agggaaactgg	ccttaggcta	atggtgtttc	70440
ttttccggaa	tgagacattt	ctcttctcac	ccatgggcca	ctaaagtctg	ctgcccattg	70500
acagaagagt	ctccaggtgg	ttgaaactcg	caatgggtag	tccagtgttg	aaattggcca	70560
tgagctcttc	aaactggccc	aagtcctatt	cctaggtgct	gcctttcttt	ttcctagtgt	70620
ttggccaggg	cacacagtgc	tctgtgtttt	tatttctctc	tctgtgtatg	gctttttgct	70680
acagtgtctt	aaagcatgtt	gaaactcttt	ttccttctcg	ttaggacaga	tatttgccct	70740
ctgcttcact	gatagacttt	aagctttgaa	tggaggcagc	aggatgtgga	gaaagccatt	70800
aggtctgcat	ggagcttccc	agggaggatt	cctcagtcct	ctcaccggcc	tctagcattc	70860
ctgtctgctt	aatcacacct	cccttggtctg	cctcagtcct	tgctctctca	actccagggc	70920
tcggcccttt	ccctgggttg	cctcttatte	cttttaagc	agtggttttc	aactagaagg	70980
gattgcaaat	ggcattttggc	agtgtttaga	gacagttttg	attgttatgg	ctgccagcat	71040
ctagtaaagg	ctaaacctac	agtgcacagg	accgcctcca	cagtggagag	acccaagtta	71100
gctatgtgaa	ggctgagaat	ccctgctttg	gagattaaaa	aaggaagctg	agggaaccac	71160
tcagttggaa	gcaccttggt	tggtcatgcac	aaggccctgg	ttctgtccct	agctctgcac	71220
aaaaaataga	atacaaggaa	gagtaacctt	aatgagctgg	tccttcaccc	agtgtgccac	71280
tgagggtcact	tgaagggaag	tctagcccca	atttagtatt	ttttgtggct	gccataacctc	71340
cagccttgat	caaactctcat	ggtatacatt	ggtaaagaaa	agggtttgaa	acatagacct	71400

gatactcggg	catggaaaca	gtatgttttg	tcagagagag	cgaaggacct	gatagacgag	71460
ggcaatatca	gagagagggc	atcagtcggg	ttagacacga	gcattccaca	gtgagcagct	71520
ctggataagc	ttttataaat	gctgggttaag	gttttgaaatt	tgcccaattt	tgtcaggatc	71580
ccagagtcta	tcacaaacat	acacagtttt	ctcaaactctg	ctttgcagta	tgcccgtgaa	71640
tgtctcttat	ctatactttc	agatggtaag	accctgaggg	cagaggaact	cagacccttt	71700
gtgccccctg	taagaccctg	ggggatgcag	tggacccgac	tttgtgttct	ctgcacagaa	71760
aggagtccac	tttcgttgag	actaagggaag	gggaactgaca	agcttccctt	tctggcttca	71820
ggttggcagt	gcctgaagct	ctgagtgcca	tctggaaggc	agaaaggatc	ctggttggtgc	71880
actgtggccc	tgagagttag	gaggtccggg	agctccggga	aatgaggtcc	tgcttactgg	71940
actcgtcatt	cgtccctgtg	gggcccttgg	tgtagagcaa	tcctcctcac	cctcaagaag	72000
gagctctggg	gatgactgag	atgttctgtt	ggcttggagc	tctcatcaga	gaggacggga	72060
ccttcccacc	tgacctgagc	ctagtgtctg	gcacagagag	cacttgaaaa	cagattgaga	72120
cactcacctg	ccatgctggc	tgctgcttgc	aagagctaac	tgccctctga	tggaaacccc	72180
atgccagaaa	aagactaaat	ccagtatcta	aaggctgctt	taaaggggtg	tactgacgc	72240
cgggcttggg	ggcacacgcc	tttaatccca	gcactcggga	ggcaggcgga	tttctgagtt	72300
caaggccagc	ctggtctaca	aagttagttc	taggacagcc	agggctacag	agaaacccctg	72360
tcttgaaaaa	ccaaaaaat	aaaaataaaa	aataagtaaa	aaataaataa	ataaataaat	72420
aaagggttgg	cactgatctg	caggcagctc	atgctagcct	aggcttttgg	ctcgatttca	72480
tctcactaaa	cgatgaatct	gtttccctgg	aacattccta	tggtttctag	tagtaatgaa	72540
gtgctgtgtt	ccactccagt	gagaacttca	attcttagtc	ttgtattata	attgaaaaat	72600
aatatatagc	aagaaatcag	tatgactgct	tacctcaaga	gacatacaat	tccacttaca	72660
atatacctgct	tccttaaat	tttcattaa	actggtgata	gataatttgg	taataggagaa	72720
ataaatcact	cttactgttg	gcagtttctt	cctgggtagg	caactctgta	ttggtttcct	72780
accagtgtct	taattcttac	tcagtggctt	tcattgagtg	ttcttggcac	tactgtcca	72840
agcactgatg	caaggcaacc	ctgtagcatg	acttcatagc	acaggcctcc	ttgttagcac	72900
acctgaaagc	agaccactct	ggctgtttca	cttgacagaca	gaatcttact	ctgtaagcca	72960
gtctagcctc	aaacaacatc	ctcctgcctc	agccttccaa	gttctaggtt	tataggaaaa	73020
ggccaccttg	cccagcttga	gactgcttct	tactgccatg	tctcttcagg	ctcacacatg	73080
aagtccaggg	cactccagga	ggagccgtga	gtctgtctgc	agggcactcc	agggggagcc	73140
atgagtctgt	ctgcagggca	ctccaggagg	agccgtgagt	ctgtctgcag	ggcactccag	73200
gggaagccgt	gagtctgtct	gcagggcact	ccagggggag	ccatgagtct	gtctgcaggg	73260
cactccaggg	ggagccatga	gtctgtctgc	aaggcattcc	aagagcagcc	atgggcgtca	73320
ctcattggta	gactgtgagg	ctacatctcc	agatgcccg	agtgcctgtg	ttgtgagcac	73380
tgctgctcat	ggtttccaac	tgagacagag	ggaaggactt	tgcccttttc	cctaaggatg	73440
ggtagtaata	gtccagacca	caagggacag	atagctatgg	ggttttctga	ctcatcctta	73500
gtacattatt	gctgatgacc	agtttggttg	gatgagttag	tgggaaagaa	gaccaagtc	73560
catacactct	gctttttaga	acttgctcat	cctagccatg	cccaaggagc	agccgttgac	73620
tgtcatggca	ttacagttag	gaaataaaca	gtcctgaagg	tgccctggcag	cagcttttca	73680
agaagctggg	gttaaaagac	agtattcaaa	catctgcgga	ctgggaactg	ggcagcattt	73740
gagtctcctg	ctgtctgtta	atttaccctg	acaaggaggt	gacttgaaag	gtttgttttg	73800
tttggggtag	agctttttca	ggaaaaaagt	ttagtcctac	agacaactct	atagttattc	73860
tagtccaaac	tcatgccttg	tgttttatct	ctaaaagccc	tgtcacactt	tgtaaaatag	73920
gtgctcttcc	tcaaaggata	tatttaacgt	tttatatctc	agggccttatt	ctgtgcattg	73980
aagctttttt	tagatgcttt	gtaagatggc	tcagtggtta	agagcatgta	ctgctcttct	74040
ggaagtccctg	ggtttgattc	tcagcagcta	acaccagctg	ttattccagt	tcctgggatc	74100
tgatgccctc	ttctggccta	tgtgagcact	gcattgtcgt	agtgcacaga	caaatgcagg	74160
caaagcactc	atacataaaa	ctaaattcaa	aaaactcttt	cattgtctca	tgtgacctag	74220
cttgagaata	cctgtgctta	tattataatc	tagtatgagc	cagccacggg	agcaacacac	74280
ctattatctc	agcactcaga	agattgagac	tagatggta	agagctagag	tctgggttac	74340
aaaacacctg	tctcaaaagt	aaaagggtcg	aaaaagtgtc	tcagcagcta	agagcacaca	74400
ctgcttctcc	agagggcctc	atttcagttc	ctaataccca	caccgagtga	ctcaaccacc	74460
tgtaactcca	ggtccatgag	atccaacacc	tctggctcgtc	tgcataagct	cctacactca	74520
attatacaga	gagagagaga	gagagagaga	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74640
nnnnnnnnnn	ntctagagtg	tttcagggtt	ttttgttttt	tttttttttt	gagacaagggt	74700
ctctctatta	tgctgcctgg	aactttctat	gtagaccagg	ctggactcaa	acttatagtg	74760
atccactact	tctgcctctc	agtactggtg	ttgaaggcat	gtgtcaccac	acccactac	74820
ttcaagatct	tagattttcca	aagaagccgt	agcctagaaa	aggtaataaa	gtactgattt	74880
aaaacagaaa	gaaatcaggt	acacttagag	ctgtagaaag	tcagcatgtg	acatttggtg	74940
caagttgtca	aaactttgct	cttaatttcta	aagagagaag	ctgtcaaaag	acttgaactg	75000
gggctgtagc	caacttggtc	gagcccttgc	atgaagctgt	gtgtttactc	cccagcactg	75060
tgggggttga	attgatttga	acccagtaga	ttcgtatatt	tgaatgttta	cctcatgggg	75120
aatgacatat	tacaaggtgt	ggcctgtgtg	gaggaattgt	caatttgggg	gtgagctttg	75180
aggtctctct	gctcaagctc	tgcccagggt	agaaagggag	cctcctcctg	gctgtctaca	75240

gaggacatag	tctcctggct	gccttcagat	caagatgtag	aactcctggc	tcctccagca	75300
ccaagtctgc	ctgcacaatg	ccatgcttcc	taccatgatg	ataatgaact	gaacctctga	75360
aactgtaagc	cagccccaat	taaagtgttg	tctttataag	agttgccttg	gtcatgggtg	75420
ctcttcataa	caataaaagc	ctaactaaaa	cacattcctg	ctgggcagtg	gtgggtgcacg	75480
cctttaatcc	cagcacttgg	gaggcagagg	caggaggatt	tctgagttcg	agggccagcct	75540
gggtctacaga	gtgagttcca	gaacagccag	ggctacacag	agaaaccctg	tctcaaaaaa	75600
aaaacaaaaa	caaacaagca	aacaaatgcc	agcatttggg	aggtagagtt	aagaagattg	75660
ggagtacaaa	gtcgtctcag	ctagtatgtt	tgaggccagc	atggaccaca	tgagacgttc	75720
tcaaaacgaa	agaaacgaat	gaatagataa	acatttgagt	gtccagtttt	ttcctttctt	75780
tcttgctttg	tttttgggcg	tgctgaggat	taaaccagag	accttggtca	tactaggcaa	75840
gcattctcca	ctgaggaaca	ccctggcgag	tgcttagtct	gtctgtctgc	ctgcctgcct	75900
gcctgcctgc	ctgcttggtt	tgtgtatgag	tggttaacctg	catgtctgtc	tgtataccac	75960
agacatgcct	ggtatctgca	gaggccagaa	gaggatgttg	gatcgccctg	aactgggatt	76020
acaaatgggt	gtaagctgcc	atgtaggtat	tcagaattga	acctgggtgct	ctgaaagagc	76080
agccagtgct	cttggtgttg	gttttatttg	gggcaggagg	tagttatttg	gttggttggt	76140
tggttggttg	gttggttggt	tttcttgaga	cagggtttct	ctatgtagcc	ttggctgtcc	76200
tggaacttgc	tctgtaggct	caaactcaga	gactctgcctg	cccctgcctc	ccgagtgtcg	76260
ggataaagtc	atgtgccacc	aactccagac	aagcagccag	tactcttaac	cactgagcca	76320
tcattccagc	ccttctttgt	gttttgagat	ggtcacaaag	tacaactcag	actgagctct	76380
tgatcacccct	ccctcagcct	cctgactgct	gggggttaca	gggtgtgtcac	tgctctcaat	76440
tctgagtgct	agatcctgaa	aaaccttct	cgtgaccttg	atccttaaaa	caaaccctgg	76500
gagaatgagt	tctgataact	atttctcact	gctcttcaag	aaaaggaaaag	ccagagaaaag	76560
gggggggggg	aagccccaga	aacattgata	acttgcccaa	agttacacag	caaaattcag	76620
acagcctgca	catctcagtg	gccatctgtg	ccatatccac	cctgcccttc	tctgacctcc	76680
ccacctccat	ccctacagac	cttgacgttg	agatcagagt	ccaagccgta	tcgtaagatg	76740
gccttaggat	ctgacatcat	ggggactctc	acgggtcctgt	cctcgtccaa	atgaaaatcc	76800
tggaggggtcg	tctttctcga	gtcaaacttg	gttaaccact	ggcctggaag	gaaacagatg	76860
gagcatcctg	agccactgtc	cccagaaagg	ccacaggtcc	acgctgtgcg	tccactggcc	76920
aaagcaacct	gagctgtcag	cagcaagaac	acaggagccg	ctgggtccca	gcatgtgtgg	76980
cacagaccat	aggctccatg	caccacgggt	tctggctatc	ctcctgtagt	aaactcagaa	77040
ataagtgggt	gttctctctc	tgacttggat	caccacgctc	cttctgttta	aagtggcctt	77100
taatagtctg	gtgtgtggta	cgtgcctgct	ctcctgtccc	ctggggactt	ggagtaggaa	77160
gccccagggc	tttctcttaa	gttaggatcc	actctgtcta	ctactccata	agatggtcac	77220
aaagcaacgt	aaaatggaaa	ttaatcaaac	catctctgcc	acaagaataa	aacagatctc	77280
aggggagggc	tgtggaaggg	tctcctgagg	ccttaccact	gtctagaagg	aagttgacag	77340
cagttcttga	gcaggggtgc	gactccagga	gttgggggct	gctctgagag	caggacagca	77400
tgtattgtag	agtgtctggg	agggagctgt	gttatcctta	ccgtgaagat	gaggacacgg	77460
gctcatgggg	gcagagccag	gattaaacct	ggctctgattc	aaaaagccag	agatctgtgc	77520
ccagccccac	gcagccattt	cactggtcaa	ctaattcaga	aacacttggt	ctgatatgct	77580
catatgctac	aagcactgtg	gccttcagat	ctccctctgg	cctggtacct	gcattcaggt	77640
tcaccaccat	caccacacac	acacacacac	acacacacac	acactcggcc	agagacaagt	77700
ggggaagccc	tcacccttga	agtaagccac	gccaaggaga	aggatgctga	gggcactggg	77760
catttccctc	gtggaccggg	caatcttccc	tttcatctgg	gcttgcaacc	agttgttaat	77820
ctcctgaagg	tctactcgag	ggttgccctg	gaggatccgg	ggcctggtcc	cataggactt	77880
ctccagaggg	gcaacaaagc	tggatttgac	tcgaagtctt	tgagaggaaa	cagatcaaag	77940
atgagagctg	aatcagcacc	ctcactttga	aagcatgccca	gaccccgagct	tcctgctcag	78000
catcttccct	tgacttgetg	gggcatctgc	cggcttgccc	agaccctggc	tagggaaacag	78060
tggattccac	cgtttgcatt	ccccgtccca	ggcctcctg	ctgtctccca	gagcccactt	78120
cctctttctg	ttcctctgtg	gtctcactgg	ctcttctcctg	cccaccagtg	ccaggcctcg	78180
cctgagcaca	cacagcctat	tgttttagaca	tcattggaagc	atacagacaa	cccaggccaa	78240
tgaagcaact	tcacgccagg	cataatgggg	cgtgcctgcc	cttcagaagc	agaggcagct	78300
ttatgagttg	ggggaccagc	tgagactcta	tagactttga	gaggtggggg	gggggtgggg	78360
ctactgactc	ctctcaaaca	caattctgga	agcactcttg	aggttcttct	caggggcagt	78420
aacagaggca	aggagctcct	tgtagggtct	gtggatgtca	gggttggtga	tcaggtcgta	78480
gtagagagcc	cggatgaatga	cagactctgt	tcgatgttca	gctcctgcca	gagagaaaag	78540
gatgccaaag	ttcataactg	cccgtagggc	ccgcatcagg	atagggagct	tagacatcaa	78600
tccttttgct	ctctgagagc	ccgaggaggc	cgatattgca	gatgttttag	ctggacaaga	78660
tcttcagggc	gtggaaaagaa	ataatgaccg	ccttgctagg	aagagctcta	agacaggcca	78720
aggttatcag	agctacagag	agaagagtgg	gatgtggtcc	tgaagtcttc	ccatcgtaac	78780
ctaccctgtt	ctgaggagga	gccagctctg	ctcacggcag	ctgtaccctt	agaacctggt	78840
taaagtacta	aaacacgata	ggaggccact	taaggaacca	aggtcgagtg	ccacttacaa	78900
agtggtaggg	attgtgtgtg	tggccccccac	cgcccccttc	ctgttctctt	gacggcggca	78960
gcatggaaac	tctgagtggg	ggaaattcag	gtccacctgc	agccttcttc	agttgacact	79020
caccagagaa	aagggcagag	agggccgtgg	ccacgctgag	tggagacagc	aggacgttgc	79080

ccgttgggct	ggcactggat	ctcaggcggt	acagatcgta	gccgaagttg	gagacagctg	79140
ctgccagctt	gttcacaggg	accttgaaga	aggggtcctc	ctcctccacg	ggctcgcccc	79200
tgctgtccgg	gactggggag	ccctgggtta	gaatacaagg	accagtaggg	aggcacagtg	79260
agtacatcac	ctcctgggtg	ggttgggtcct	ctagtccctg	gggccatgag	tctgaggtca	79320
gaatgagtgt	gtgctctctg	actccacaac	ctgtgtgctg	ggaggtgggg	agtgggaagg	79380
gcaacacaaa	agggtctg	agacctgaac	tggtgtctga	gaacctgaag	cctggcccac	79440
tttaaaataa	aaactgtagg	gctggggaga	tagcacagta	gataaagtac	cagcatgcaa	79500
gttcaaggac	ctgggttcag	tccccagagc	tgggcacggg	ggtgcatgct	tataatccca	79560
acactgggga	ggcagagatg	ggcagggtcct	ggggctcatt	ggccaatcag	cctgaactaa	79620
tcagcgtatt	ccatctcagt	gaggggtcct	gtttcagagg	gcctgaggaa	tgactctggg	79680
ttgactacta	gcctactctg	tgtctgtttc	tgtctgtctg	tctgtctgtc	tgtctgtctc	79740
caccctcttc	tgtcccttc	cctctgcagg	gaacttcctc	accaccacca	acccccaaag	79800
aaaccacccc	tcagaccagt	cttccctatt	cagcttgctg	gctgggtccta	gtctgcctag	79860
gttctgctgt	gacgccctcc	ctgtctttcc	tgacaagcca	tccctctga	ctagaccoga	79920
gaggaatttg	tcgttttctg	acctgttttc	agtgtcagcc	tcttccttat	gagactttct	79980
gcttttttgt	tttgttccca	gggctttgga	tcaaagctgg	gctcttacat	acgttaggca	80040
aatgcttggc	caccagctg	tacctccgt	cctgtgtgt	tttcggtttg	gaggactttt	80100
tttttagttt	tctgttttgt	ttttgtttt	gattttttgt	tgtttgtttg	tttgttttga	80160
gacaggattt	cgctatgtga	ctctagctgt	cctgggactc	actatgtaga	ccaggctggc	80220
cttagattca	gagatccacc	tgcctctgcc	tcctaagtgc	tgggattttt	agattttaat	80280
ctgtacctac	caacctcaaa	ggaagtgtcc	atggatagag	ttcagtacta	catcatgtgt	80340
gtaacatgtg	tgagggcctg	ggcttcaccc	ccaacagaga	agggggagtt	agtaggtgag	80400
gaataagtg	ctggctagt	gcaagacagt	attgtctaa	gtcactaagc	cttaagccac	80460
acttaaagcc	cacaatccag	gtctaata	cccatctgcc	ttgtccttgt	gtgacatgac	80520
cccaccccta	cttcctccgt	atagtggcag	ctcctctgga	tctgaaagg	agaggggaaga	80580
tattcttgtc	tcgatgttaa	agtaaccaag	gcctagaaga	gtgaaggcca	aagcccaccc	80640
tggatccagg	gctgctctcc	tgcaactgtc	cttctgctgt	cccactacc	caccctactg	80700
acctcagagc	tgctggggac	gttctggctg	ctgccgtgcc	cgagcagggc	tccagtccag	80760
aggagtagca	ccagggcctg	catcccggaa	ctacaagaga	aacaagagag	caaacgactc	80820
ccctcaccca	caccctcccc	tgccactgca	cattgcacac	tgacacagga	caagagtcag	80880
gcagagttag	cccttcccc	ccctccagct	ctcagcccca	agtggaccct	tgacttgagg	80940
tcttccgtcc	ctgacctgcc	cctgcacttc	tccttagctg	tgcccccat	gttggttcct	81000
atcaggagac	ccaccttccc	atctaagctc	cagcacagg	aagaccagc	agcaggctct	81060
tcaggcccca	agacaatgct	ctagcacaaa	cacacaccaa	ggcttttccg	tggaggcaca	81120
cgccagctc	ctttggtagg	atttggaacc	ctgtctcagg	atgggagcag	agcccagggtc	81180
atagacttac	agaacatctg	gtctggtcct	gctatccacc	aatagttctc	tgaccaaaagc	81240
ctatgttaaa	gacacacaca	cactttcttc	taggttaggt	cttgtgtatg	tagcccaggc	81300
tagccttgaa	gttgacgcta	ccctacttca	tttgctcctc	gagtactaca	atgccagggtg	81360
tgagccatca	tgctgactt	gactcccttc	ttctatttca	agagcaattc	ttagttaaga	81420
gggtatgaac	cagggccaca	ctgccnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81540
nnnnntctta	tttttagctt	gtttgttttt	cttacttgag	acctggggag	gggaggtgta	81600
tgtgtcttcc	atttgcttct	tctacctaat	aaagtttctt	gggtgatgtt	gggggggggg	81660
gagggggtag	gactcaagaa	gggattctcc	cataagctgt	tccgtttggg	tagtactatg	81720
taagggaagt	acaggtgggc	agagctcggc	tctgcctgac	tggcgtgctc	tgaggtaaag	81780
gtgagatggt	gcaagatttg	ggccctcagg	agttggctct	ggtggccctg	taccttcttg	81840
tctgtgggta	aggatgacca	gtaggtgaga	gatgagggaa	ccagaacaga	aggtgaaagt	81900
tagtggggcg	gagccccaga	ctagttaggt	gggggttaaac	tagatgactt	tctggaaccc	81960
caaggggctc	ggagactagt	gggtgtggag	aagacctcta	atgtgttgta	aggcctctga	82020
actcagtagc	cgaacttgat	gccagaaagc	cccaaactgc	taaacccaag	caggagcggg	82080
acgccatccg	ttccatggct	tcacccgagg	tggcccccag	gctgcgccaa	tcaatgagca	82140
gccgagagat	aggggctgtg	acaagccagg	aaaagttaca	gcacgctgga	aagataatac	82200
aggccaggaa	gccccaggca	cagcagggtg	gaaaagctag	atcccgatcc	tgccggaggg	82260
gggccccttc	gaggtcccgg	gcacccgggtg	caaggatcag	agaaactgac	tgaaacctag	82320
ctgacctgcc	cagacctagg	catcctgggg	actccttggt	gctggcgctt	ccttcacggc	82380
gtttcgggga	ctgcaactgg	ggctgcagct	gctgcccacc	ccgaaatctg	ttcgggaccg	82440
ctggatgtgg	cggaacattt	tcgtttcgct	gatacacagc	ctactctctg	gagtaggggc	82500
gctggtcggg	tgcggaactt	ggggactgac	aaagcactga	ggggcggggg	tggaagagag	82560
ggcctggaag	actgaagtgt	gaaccttttg	gaatgggaact	ggtttggtgt	gtggatgggt	82620
gggagtaacc	agtgggagaa	tggatctagg	tctgggagaa	attgacctta	gctctttgtc	82680
ttctccaggc	tgtggcagtt	tcctcaaatg	gtcaccgacc	caattaatga	tcacccaccg	82740
tgggcacggg	tcctagtagc	agtgtcagtg	ggtgagtgtg	cagaaaaggc	tgaatcggga	82800
aaggccttgt	tggaccggga	attctaggtt	cctcccccat	ctttggaatg	gagcagatgt	82860
tgctggaggt	ttgctgtgag	gaattaagga	cctgagaaaa	gtgggacttg	agatatctag	82920

gctgtgcatc	agctctgagc	gaggagcctc	atagtcttct	ccggtgcctt	cagggttattt	82980
cgctgcagat	ggagttgata	tgctgtggaa	ccagacattg	gccagggcct	gggaccttct	83040
ctgtcaccat	ttggcggtaa	gactctgaag	ggagaggcca	ggtagtaagg	gagcatgtcc	83100
aactcaagg	cccaacctct	ctcttcagtg	ttctgtcctc	tgacttttcc	acaaagcccc	83160
ctgaaaacct	atcctctcag	acttggattg	agttggagg	aggttttgac	tggctagcca	83220
ctcctgggca	ctgcccagg	agtttggttc	tccccacaaa	cctccagctg	atcataaaaa	83280
aaaaaaaaaa	aaagccagga	atgaaagcta	gggtatgcta	tgcaaatagt	gtggccttgg	83340
gtaagagaac	ctctggtcca	gggctgctca	tgccccctag	ataagggtca	gcagaaaagg	83400
caggatttga	ggcagtccta	aaaaatgctt	gggtaataata	aagtgaataa	ataaaaaata	83460
aataaatact	aattttttaa	aagctgatac	ctggaaggat	gaggcagaga	gtagaaaaaa	83520
catgcgtggg	tgtccctagg	ataaggagct	gggacttggt	gggcacaggt	catgcaaagc	83580
ctgaaccttg	aaccttgctc	gcaggtagtg	agctgcctca	gcaccgctgt	tgtgtctggc	83640
cactatgtgg	gcttctctat	ggtatccctg	cttctggagc	tgaactccat	ctgtttgcat	83700
ctacggaagc	tactgtctgt	ctcccataag	gccccatcct	tgcccttcag	agtaagcagt	83760
tgggcccagc	tggccaccct	ggtcctcttc	cgccttctcg	ctctgggatg	gatgagtctg	83820
tgggtgtccc	ggcagcacta	ccagctgtct	cttgcctctg	ttctgctttg	tgtggctggg	83880
ctggtccacc	tgggcagcat	aagcatctcc	acagggatcc	gaattctgac	caaggatatc	83940
ttgcagtctc	agccctaccc	gtttatcctc	atgcacaagg	aaaccaagac	acgtgagcct	84000
gttgccagga	acacttccac	tctcagctct	aaagggtgtg	aagttttctc	ttctgtcagc	84060
ccccagggag	gtggggctgg	gaagaggaga	tggtagccca	ctgcatagtc	tactatgtag	84120
caaggactag	actgtatcat	cagagagaga	gagagagaga	gagagagaga	gagagagaga	84180
gagagagaga	gaacattgta	tgagatctcc	attacagctca	ggaaatcagg	agatctaaat	84240
aacttttaaaa	gtcccacagt	ctttacatat	tcttaaaatt	tcaatctctt	taaaatatcc	84300
atctctttta	aaattcaaag	tctttttaca	attaaaagtc	tcaactgtgg	gctccactaa	84360
aacagtttct	tccttcaaga	gggaaaatat	cagggcacag	tcacaatcaa	aagcaaaagt	84420
caatctccaa	ccgtccaatg	tctgggatac	aactcacgat	cttctgggct	cctccaagg	84480
cttgggtcac	ttctccagcc	aggccctttg	tagcacacgc	gtcatcctct	aggctccaga	84540
tacctgtact	ccactgctgc	tgctgtctct	ggtggtcatc	tcatggtact	ggcatctcca	84600
aaacgctgca	tgacccttct	agtcctgggc	cttcaagaga	gaagactaga	gcctggcaaa	84660
gtggcacatg	ctgataatgc	tagcacttgg	gaatgacaag	cagaaggatc	agaagttcaa	84720
ggccagcctg	ggctacaaga	gactctgttt	caacaaacaa	acaaacaaac	aaaccaaaga	84780
agagaaagaa	aaaactggac	atgacagccg	gaacattatc	tgacattcat	aaggtcctga	84840
gttcaatgcc	aagttggcag	tgccctagtt	gataagggtc	tagccactct	ggtaatacca	84900
tggctgactg	aacaccttac	ccagcaactt	gctgatagac	tctgccttcc	agcaaaaagg	84960
aggagcttgg	ctgaggagag	aacattgaac	cctattgtat	atgaataaat	tgtgtgtcaa	85020
atgatttcat	cagtctcttg	tgaatgtgat	tgctttgagt	catttttctt	ggctccagtg	85080
ttatcctggg	ctgcagtgtg	gtgtggagtt	gtggaaagct	tgagttggga	gggtttcctg	85140
ttaaagtttc	tctggctctt	ttctttcttc	ccggtttttg	ttttgtttgc	ctgggtgggt	85200
tctctgggtg	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	tagaagttgg	85260
cggggggtgg	agggggctgg	agagatggct	cagcggttaa	gagcgccaac	tgctcttcca	85320
aaggctcctga	gttcaaatac	caacaaccac	atgggtggct	acaacctcc	gtaacaaaaa	85380
aatctgatgc	cctcttcttg	agtgtctgaa	aacagctaca	gtgtgcttac	atataataaa	85440
taaataaata	ttaaaaaaaa	aagaagttgg	catggatgat	gtagtgaaga	ctggcattag	85500
atatctcttg	atccccctgc	ctctacctct	tagacactgt	gagtatggaa	gtgtaccacc	85560
gcaccaggcc	aggctagaac	attctctgat	ctacaaatac	ctagagtatt	attcctctat	85620
gatcagaaaa	cagacccagg	gggcccacaga	aatgtcttag	taggtaaaaa	cacttgcttt	85680
caggcctgat	aacctgcggt	tttttgtttg	ttctgggggg	cgggagaggc	tggctggctg	85740
gctggcctgg	aattccacaga	gatccacctg	cctctgcctc	ctgagtgtca	ggtaccagga	85800
tcacaggtgt	gtgccaccac	acttggecta	actgcctgag	tttgagcatc	agtactcaca	85860
tggtagctag	gatagaatag	actctcacca	gctcttctga	cttccacatg	tgccctgcag	85920
catgggctct	ccttccccaa	aggaaaaata	aatgtaagaa	ttaaaaaaaa	aaaaaaaaag	85980
caaaccagg	tcttgtgtga	tggctcagca	tcaaagctac	ctcccgccac	agctgaccac	86040
ctggtgataa	cttatagcct	tggtatgctc	tcccttgacc	cccacgggca	tgctgtacac	86100
gtatgtgtgc	ccacacaaaac	acacaatcaa	gaaataaatg	cagccaggcg	aggtggcaca	86160
cccctttaat	cccagcactt	gggaggcaga	ggcagggtga	attctgagtt	cgaggccaac	86220
ctggtctaca	aagtgagttc	caggacagcc	agagctacac	agagaaaccc	tgtttcgaaa	86280
taacccaaaa	aaaaccactt	taaatattat	ttttattttg	ttttgtttat	cctggaactt	86340
ggtctgcaga	ccaggctggc	cttgaactca	cagagatcca	actgcttctg	cttcccagc	86400
acattaaagg	atgtcccacc	actgcctggc	taaagattta	ttttttcttt	ctttttgttt	86460
tgttttgttt	tgttttttct	aaaaaatttt	tttaaaaaga	accatccctc	ctagcactca	86520
ggagactctg	aagtcagggc	cagccaggtc	tactgagtga	gctctagggc	agccagggct	86580
ccacaaagaa	accctctctc	aacaaacaaa	caaaagagaa	cagacccaac	cagacctgag	86640
gacacacact	tgtaatctaa	gcccttgaga	ggctgagaag	ttcaaggcta	gccacaagtg	86700
tgtggtgcat	tcaagagcag	cctgggtggg	ctacagaaaa	agaaagagg	agagagagaa	86760

tgggttaatga	agatgactct	ggaaaagtga	aactcaagag	aaagcccctc	agattttgctt	86820
aagacgagtt	gaggggtggag	aaccgcca	gcggacgagc	cagacagaga	ctgccaaaca	86880
agttcaatcg	gttcaggtac	attacttcca	aaacgccatt	gccacatcag	gatgcttcaa	86940
tcagccaaac	caacgcagcg	actattgact	tctgcatttc	agagacttcc	gtctctgtcc	87000
agggcaatgt	cacttttagct	ttcctttgca	gaaaggaaaa	gtccctgcct	ctgatgtggt	87060
agatcctcac	acaccttctg	ccagatccag	acactgggtat	gactcagcct	cggggagctc	87120
tatctacaga	gataagggtg	caaggcgtgt	gtgttttaag	tatgtgttta	aaagtacaaa	87180
gtgagagtcc	ctggaaaggg	ctccctgccc	tcaccatcac	cgaaagcaca	aaccttaggg	87240
taatatctga	cattcctgga	aatgtatytg	tgtattcatt	atgtagccct	gactgtcctg	87300
gaatggggta	taaaccagga	tggtctcaca	tctcagagac	ccatttgcct	ctgcctccca	87360
agaactaaga	ttagaggcat	gcactaccat	acttggctca	tgatttactt	aactttattt	87420
tatgtttcacg	aatgttagcc	tgcatgtatg	tgtgtgcacc	atgtgcatgc	ctgggtgccc	87480
agagccacag	agaagggtgt	gggttgattt	cctggagatg	aagtcccaaa	caactgtaag	87540
cagtccaatg	tgtgtgctgg	agatgaaact	tggttcatcc	acaagagcag	tatgtgctct	87600
taactgtgga	ggcatatctc	cagcctcaga	tttcccagtt	aatgtttgct	ttcgcaccca	87660
ggcccctctg	cgcgtgcgct	ggagacctcc	tttaccgcct	tgagcctcat	tgggcaattg	87720
tggtctgggag	acttgcagat	cccaagtgtg	acaagagaag	aataaactgg	tgtgctatga	87780
actcactctc	tctctgtagc	cattggctga	gcatactttg	cctcaaccta	ccgcccttcc	87840
ttccccta	cctaaatctt	tgcctctctc	aaatgtgctc	ctcccccgca	gtaatccagt	87900
ggctgctggg	gctctagaga	gatggggggg	gggggagcaa	cggttacagc	ttaaggcagc	87960
tgacgagaa	cttttttgct	gtatattgag	tcttaaaaa	tcatataaac	tttggttctt	88020
gtttctaaat	ataaccccat	ctgtttcaac	acaaaatgca	acaacaaaat	gtttcaaatt	88080
gctatttgga	ataattaaaa	aatttcaata	cttgatttaa	aaatgcttta	actttttaaa	88140
taaattttaa	atgttattat	ttttaaaaag	ttacaagttt	aaaaaaaaga	aagatagaaa	88200
tcacataatg	aaattaacca	tacgcaagtg	aggctcggtg	cactgggtaca	cagttacagt	88260
agccatttgg	agtggaggcc	atggcgcttc	acattgaatt	ttatactttc	tttatagaat	88320
attttttatg	cacctatcta	ctactgataa	caaaaacccc	atgagagagt	tagaattaga	88380
catcaattag	ctttgatcct	ctgtcataac	tcgtgtccac	tccctgcctt	agtcctacct	88440
catccctgtc	ctcttttcta	catcttatac	tgaatccaca	cactcagttg	tttacacaaa	88500
cacatacatc	actgtccann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88620
tgccctggctc	tgttttgctg	ttgttgtttt	tgtttgtgtg	tgtgtttaga	tttggtttgg	88680
tttggtttga	ttttgttttt	ttttagagaa	tcttactatg	tagctcaggc	tgtccttgaa	88740
ctcacagaga	tctccttgct	ctgccttcca	actgtctgag	ttaaagggtat	acaccacctt	88800
acctggcccc	tttcatctat	ctatctatct	atctatctat	ctatctatct	atctatctat	88860
tatctatcta	tctaaaattt	atctgtgtgt	gtctgtgtgt	acattcccca	gagcctgtgt	88920
ggtagtcaat	aagtaaccct	cagaagtgtg	ctttctctaa	tcttggatc	aaacttgaat	88980
tgtaggctt	ggtagcaagc	atgtttacc	actgagccat	ttatgacccc	atggccagc	89040
atcttccatg	ggttctggg	acacaaatgt	gtactttgat	gtttacagga	caagcgctta	89100
accaaccaag	tcattttccc	agccccatcc	tgactcccat	taagtgttct	ttcccccaac	89160
ccaggaccaa	atctagagga	gtgtccatgc	ccaacaaaaca	ctctgccaa	cctctccccct	89220
tactgctctt	ctcccttccc	ttccttcatt	tcttctgtcc	ttcttttctt	tctttttgaa	89280
acaggtcttt	tctctgcate	ccaagctagc	cttgaaactg	tgatgtagct	caggctggct	89340
ttgaactcac	agctgtcctc	ctacttcagc	ttcccaacaa	ctgggattat	agacctatgc	89400
taccacacct	ggctcatttt	tcaaataaat	aaaaagaaaa	tcaaaaagtt	cctagaacag	89460
tcacaggatt	cacaaaaact	ttggaaggag	actaaaaatg	gattttttaa	aatgcttga	89520
agcaciaaaga	gttggtgaaa	gaagagagaa	gaggaaaagt	tagcttagta	ggtagaagtc	89580
aatcaagcct	cacaccctga	gttcaattcc	tgaccctatg	gtagaaggag	aagagcaatg	89640
cgggaaacat	tatcctctga	cctccagacc	cgctgtggca	cgtgcatgca	cacacacaag	89700
caggagccct	tggaggggaag	tcctagaaat	gaatcttact	gaagcaggtc	tgccaggccc	89760
tgtgctcagc	catttttattt	ttcctttgtg	taccgcacac	gcttccattc	tcaaagttgt	89820
gagtctgaga	ggaagtactc	actgtgtccc	cagttagctt	ctgtcttacc	ctgggtcact	89880
tagatgggggt	cacttagtggt	tagccttggt	gtggagaaag	agaacacagg	tccgagtagc	89940
cagtagacct	gagtctttat	atctgcaaag	gggtgtgggg	cataatcaaa	tctccccccc	90000
tccccgggggt	cctgatacca	ggttgattta	agtgtatgtg	catggctcctt	ccaagtcttg	90060
acacatcatc	caactccaag	tggctttctc	atttttcctt	gccagtagcc	tcttggtgag	90120
gaaatggctg	aggaaaacag	agttgcagaa	agacagggcc	atggcctggc	tgaggctttt	90180
ctctgagtct	gaagagggtc	agcgactctg	agaaatgaag	ctatttctga	gtgagagggg	90240
ccaaagaagg	aacacggcag	agggagagcc	cccagggaga	tggagacaga	agccggagag	90300
ggacctgtgt	cgaggctgga	ggggagggaag	agggggggag	agttagaccc	actgtcatct	90360
gttgggcaga	gaggggctac	attcatctgc	agtatgggtg	agaggggaca	gagagtgtat	90420
gtaacaggaa	aaatttgggg	ttgagggggg	cagcctgtag	ggctggggcc	cagcagtgtg	90480
cagctaggta	gagtagacag	taactcccag	aattctctgt	ctccactaaa	tccctgttcc	90540
gctccgtgca	gagtaaaacc	cacacagggt	ggatttcagt	ctcctttgca	ccccctcca	90600

ccccccctcc	acccccagct	ctgggtcacag	ccagtcagag	ttgggggtgg	ggcagatctt	90660
gtaaaagagg	ctgggtgagga	catcggaag	tctgtaccct	ccactagcaa	agtgccagac	90720
gctccgtgac	acttttaaatg	cctcagataa	aacagtgaga	gactctcctg	gtggcaggca	90780
agatgatggg	tcaggggacct	cagcgccctct	gaggctcaga	caccaggata	aagaataaaa	90840
acaccacgga	gacccctgtg	acccctcgt	gcagaggag	aatgccaatg	tggcccagct	90900
agctgctgga	gcgcaagcct	caaggccctgt	gctagttagt	agtctactgc	tgctgccttg	90960
tcccaataaa	cccccttccg	cccaggatta	gtggacacgc	cttgctcaag	ccgagtcctt	91020
gatctgccac	cttatcacac	acatacaaaa	atcccttgag	gatggttacc	atcctgggac	91080
aaagcctcat	tctctgctta	acccaagtga	cacctatatg	gcagatccct	gtgtccttct	91140
cctgatgata	acaacccttg	caatccaata	gaggggaact	cgggtttctg	tcagcttccct	91200
ttatgctgat	agaaatgtac	tctgcatgtg	gggagcctgc	cttgctcacc	ctgagacccc	91260
atggggctgg	ctggggcttt	gcacatcatt	gggactcaga	gatggtgact	acatgaacgt	91320
cccacacttg	gttgcacaa	gcagaatgac	aggatgttat	gcctgggtgtg	tgagtgtgtg	91380
tgtgtctctg	tgtgtgtaaa	acgcctctct	ctggagccct	cctgtctgtc	tgccctctgt	91440
tcaatggctg	cacaattgtc	ctttctcttt	ccaaggacct	ctgtatgggt	gtgtccttca	91500
ttcagtgcct	ttcctctgtg	ggtttgtcct	gtagccccc	tgctactgag	aaagtcttct	91560
gtctgtcctt	gggttgtctg	gctagaacac	agacatcatt	gtcttttttt	tttttttttt	91620
ttttttttaa	agatttattt	atatgttaagt	acactgtagc	tgtcttcaga	cactccagaa	91680
gagggagtca	gatctcgtta	ggatggttgt	gagccaccat	gtggttgctg	ggatttgaac	91740
tccagacctt	cggaagagca	gtcgggtgct	cttactcact	gagccatctc	accagccccc	91800
acatcattgt	cttgcccacg	actgtctctc	agaatgggtg	gcaggaggat	gtgaccccc	91860
acccccaggc	acggggacac	aacatcttct	acacgtgtag	gtcttgtgca	ctggccttgc	91920
tttcttcttc	caagcaggtc	tcccaggaaa	tggcacttac	agagattgaa	gagtttaata	91980
catgtctcgc	tgccctctctt	ttcggaacc	cccagagggt	agcagcagaa	accagggctg	92040
gcaggggctc	taagctgcct	gggcaaagga	gcagggggta	gcatggagcc	ttagccaatt	92100
tggaaagcac	tgtgacccaa	gcacattttg	cagcagtaat	gtcaaattct	gccgttcagg	92160
catgccattg	atgtgcacgc	tgccacacag	aaaccagtga	cacaaaggca	cagccttctc	92220
caccctcctg	gtgcttagga	actaacggct	ctaatagaga	atgagagctg	aaaggagaga	92280
gacgggggcg	ggccacagca	gcgcaggctg	gcactgcgtg	ttggaggagg	ctgacccact	92340
tctcgtagag	gtaaggggcc	cactgaaatg	tcacttaaat	tagccaccac	tcccaacact	92400
agatctcctt	tgtcccccata	cctcagcccc	acgttctttt	ctttttttct	tctttttctt	92460
ctcctctggg	gcagcctcaa	gcccagcacc	cactttttag	agctgtaaac	caccctggtc	92520
ctagaagccc	tcttacgtta	ggggatgaca	ggaggtagag	atcaggaagg	agggaggagg	92580
gggaggagga	aaggaaaagg	gaggggagag	agggaaagag	atcgagagag	catgcattca	92640
tcacaaagag	ccctctttttc	tggctttttg	actgcactgt	gagttattta	gccaacaata	92700
gatgtttatg	tatttttttta	gaaccctgat	ttattaacag	cctgaaaagga	gagagacgga	92760
gatttatata	ggaagtgcag	tgagttaagg	ggggcaatta	agagagcaga	aagagatacg	92820
gaacacagac	ttgtaaagg	ttttgtaaca	tccaatcaaa	ggtgcttcag	gtattttcca	92880
aggaagcaga	aggtaaaaaa	aaaaaaaaat	tgtcccata	gaagctgaca	ctggatggag	92940
caatggccca	ggcggaactc	ctgcttgaaa	gaaggtgaga	agggaggggac	acagaccagg	93000
atccgatgag	ccagagtgtg	gccatagctg	ggtcatgagg	cccagggttg	gaaggacccc	93060
actaaagtgt	gcactggcct	ttccttgaca	aaggatgcac	ctatagctag	gcgtggtggc	93120
aagtggttgt	tattctagta	cttaggaggc	tgaggcagga	ggatcaccat	gagtgtatgc	93180
ccagcctgga	ctgcatagca	acacccagtt	tcaaaataac	aacaaaagga	agtgggggtg	93240
gggagggcaa	cattttggaat	gtaaataaat	aaaacatttt	ttttaaaaaa	agaaaggggc	93300
tagtgagtta	gttcagcggt	taagagcgct	gactgtctct	ccgaagggtc	tgagttcaaa	93360
tcccaacaac	cacatggtgg	ctcacaacca	tccataagga	gatctacgcc	ctcttctggt	93420
gtgtttaaag	tcagctacaa	tgtacttaca	tataataata	aataaattct	ggagtggagg	93480
ggccagagca	agtagaggtc	ctgagtttaa	ttcccagcaa	ccacatgatg	gctcacaacc	93540
atctgtacaa	ttacagtgca	ctcatataca	taaaataaat	aaataaatct	ttaaaaaaag	93600
aagaaagagg	gtggggcagg	ggagggaaga	agaagaaagg	taagaagcta	aataaaaaggc	93660
acagagatga	gcttcatgtg	gaaacacagg	cctgtagtcc	tggcactcag	gtgggggttg	93720
ggggggctac	agtgagagta	tcagttagttc	aaggtcaact	tgggtgagac	cttgtctcaa	93780
aaaatacata	ngcnaaaaaa	aaaaaaaaaa	acatagccag	gcatgatggt	atacatttat	93840
agtcccagca	cttagaggac	tgaggcaggg	cagaaagaaa	aggaattcaa	gatcaggctg	93900
agctgtatgc	agtcctgatc	ctatcccctc	ccccccccc	ccagagacag	acagacagac	93960
agacagacag	agagaaacac	aaagaaaggg	gccttcagat	ggctcagcaa	ttaaaggcgc	94020
ttgctattca	gaccccatga	cctgagctca	aagcctggga	cccaaggtag	aaggcaagag	94080
ccaaactccac	agagctgttc	tatgatctct	atatgaatgc	tggggcatgt	gcctacacta	94140
tggtgtgcac	acatgcacag	attagaaaaa	gaggagggaag	aaaaacataa	gattgtttca	94200
agaaaagaaa	ggctggcttc	ttccacgtca	gtgtgagagg	aggggtctggc	ccctttgtag	94260
ccaggtcctt	cccagtcacg	tgggggctga	actgaggcag	cggaggaggc	aataacggag	94320
ctttcccaac	gcagtgtcca	gcaaaactcaa	ctctacagcc	tgtcctgatc	cacagagaag	94380
ccttctctggc	tccctcacca	atgcgggggc	attggctccc	aggctcctgg	gccccccccc	94440

acacctgtgg	agtgtctaggt	gatttgctaa	tgttgggcaa	catttgccca	cgtaggggttc	94500
ttggctcttt	ggtaatatagac	atgcctagca	ggagggcgga	gcttggagg	gggagtcctg	94560
gggttgcccg	tggctccctg	cagctgggg	gtctggccag	ctgaagaagg	agccatggca	94620
cgcaaattggg	agagcatgga	acagaggctg	tggatgctaa	gcaatatggg	aggcagtcta	94680
agcttggaag	cagcaggtgt	ctgggaacgg	gctgtggcc	caggcagatt	tccagtggagc	94740
actccagttt	tttggcacia	ggaacaagct	ggctgagccc	aagaggcaag	tgggtgataat	94800
gaaacccgca	ggtgaggaac	agcgggtaag	ggtgccatgg	gagcccatgt	gctcatgaag	94860
aggctgggg	gtgaagaaga	gcccatgcag	ggaagccaca	catccccctg	agttccaggc	94920
agaggcagag	tccctgagtg	gggctccctg	ggtctccctt	tacctaacca	gtctcccgcc	94980
accccagcaa	acaaaatccc	atccataatt	tgagttttat	agagaccta	aaggctgagc	95040
tactgtgtgc	cactaacat	cagcctaacc	ctccccact	gtcttctcta	gctgccccctc	95100
tttcttctga	gactgtgata	gtggcgggga	cgggttggga	gtgtgtgtga	agccctctcc	95160
gactctccaa	ccccagctga	gccccctgtt	ctgcagctca	gtaacacagt	aacacaggct	95220
cagttctaca	ctgggttgaga	acactcacgg	ctctctcagc	tccttagaga	gcctgttttc	95280
tcattttcct	gtccccaaag	cctagacaat	ggctgggtcca	tttgtaagct	tatctgagga	95340
tgccaggggc	caccccatgt	ctccactagg	ctggcaatgt	tctctgtcac	tgtagtacag	95400
aagactgcct	ggtgggaggt	gagataagga	aagggtatgt	ctccccctggg	gttccccacac	95460
agtgtctgagc	ggaaaatggc	agaatgggct	gggaggtaac	tctgttgcta	gagtacttgc	95520
ctagcatgtg	caagggaagg	cctgggttcc	atccccagca	ctacagaacc	caggcgtggt	95580
ggttcatgct	ggtattctca	acattcagga	ggtacagtca	ggaagagcag	aagttcaagg	95640
ccatcctcag	ctacatagct	agcttgagac	cagcctgggc	tatgtgagac	tttgtctcca	95700
acaaacaaca	acaaagcagc	agaaggccaa	ctggcaagag	gagtattacg	taaagtaa	95760
ccatctcaaa	aagcaagtag	catgtatctt	ctttcatttt	tttttacatt	ctataaaggc	95820
ctatcaagtc	atgtatatat	gcattgtatgt	ttgtatgata	tgaaagtagg	gggctggata	95880
gatggctcag	cagttgagag	cacttggatg	ctctttcaaa	gaacctgggt	tcaattccta	95940
gcacccacat	ggcagctcac	aactgtctgt	aattccagtc	tcaggggatc	tggcacctc	96000
acacagatat	ccacgcacat	aaacaccaat	gcacataaaa	taaataattt	ttaaaaaaag	96060
aaattggaag	taaaactctc	taaggagaca	aaagggaactg	agggggaagt	ggaggggcat	96120
gaagggggag	ggcataggtg	tgtgggtgtg	ttaacatgca	gaatacactt	ctataaaaagc	96180
tttgggggttc	attatgcaat	gtatacatgt	gtgggtgcaa	gatgtaagct	gtgcatatgt	96240
gtgggggcca	aaggctctct	cctcaatccc	tctctgcctt	attttcattt	aaattataat	96300
tattactatt	agtgtgtggt	gtgatgtgtg	tgggtgtgtt	aagccctcac	ggcaatcaga	96360
ggatgtctgt	ggtctgagga	tgctctctta	ccatgttctg	gtgggttctg	tggatggaac	96420
tctggtagtc	aggtttgcaa	agctagtgtc	tttatctgcc	gagccacctt	gctggccttc	96480
aaccttattt	tttgcatgga	acatggaact	tcctgagttg	cctggacagc	aagtccccaa	96540
gacctctctg	ttcctgcctc	ccccntgtcn	nnntcacang	aggacacacn	gcttantggg	96600
tntccggatt	gctgcncacc	tccccgcenc	ccnagcctcc	tgccctccccg	ccccctgcgc	96660
ccgctggnc	ctcccccccc	cccccccccc	cccccccccc	ctcccccccc	ccccnnnnnn	96720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96900
agctctaact	gctgttaaat	tcacactcct	tcacatcccc	acgctaggac	tctaaggagg	96960
caccagcaag	gcccagggtcc	agcttgactt	agagcaaagc	atccctcccc	ctccacacaa	97020
tggaaacgga	cggaaagggg	catggaagca	gaaccagaca	acagcagcct	agccaagccc	97080
aggactctgc	tccttcccc	catgcctgcc	gtgcaactgg	ggaggcaaag	ccccagccgg	97140
tgctttctga	ccgcttagcg	gaagacaagg	ggagcctgtg	attatgattt	ctgctgattt	97200
gcaatgaaac	actaatgcag	tgggcttttc	attaagccag	atttattcaa	tctaaagatt	97260
ttatttctct	tatgtagaaa	gtgcatcttt	atatgttgtt	ggaggagcag	agatgtgata	97320
aaaagaaatt	tctcttatga	actaatagca	ctgatacata	gtggtagcta	tgccataggcc	97380
tctctctctc	tctctctctc	tgtctcctgt	gcattgtgtg	gtgtgtgtgt	gtgtgtgtgt	97440
atgaatgcac	acaaagtagc	ccccccccat	attatttctt	ctgtgggata	tccagactca	97500
gcaaatgggt	gtgactggga	agtctggcca	tgcaattctt	gccttttctc	ttgccagccc	97560
aatccctttg	cattcaaacc	cgggctgctt	gctgtggcca	gcccctttcac	ctggagtcc	97620
tctctctctc	tacctgtctt	cccatccttt	gcagacaatt	atccctcaata	actagccaat	97680
tacccttaag	gacaattata	ctcttccatc	agcaaacacg	ggtgttcttt	ccttgagtct	97740
tttgatgaag	tcgatattaa	agagatgctt	tatttacata	aagtcaaata	gctccctttt	97800
agaagggttt	gggttcgatg	tcaaagtttt	aaaatcttaa	ctagaggatg	ggtgtagagg	97860
gcttttggct	agggtagaaa	agagatggag	atacttattc	tgatgtgtgt	ttaaaaggta	97920
ggatgccag	agaagggtga	aggatggggg	aggaggggtc	cctcctcaag	ctaagtaac	97980
taaaagcagg	gatgagctgg	gcgctaggag	tggaaaggag	cagaagtgtc	tgcccttgac	98040
tgaccacagc	tcctgccttc	ccctccccca	gtctctctgt	gaaccgcccag	cattaggagc	98100
taatcgcttc	agaaagccag	attggaatgt	gttgctcacc	ctccactgct	cagaaaacct	98160
ttattccagg	caaggactga	cccaaaccga	tcattggcatc	tgccaatcag	gaggccaaaag	98220
gtgcccggcag	ggcgggacct	agctgtgcag	aaacagctcc	gttatggcgc	gcagaaaaag	98280
ctgggggggaa	aggctaccgt	tttatctctt	ggcagatggc	ttctctcttt	gatgctttgg	

gccttacctg	ttactgcctg	cacttgactt	gacctaggca	aaaatagcag	cgagatacac	98340
gttctcgaag	ttagaaggaa	aaaaaaaaag	ccccaaacca	caacacaacc	cggaagtgtg	98400
ccccgcgtgt	gtttctaaag	agctgttttc	ttcccaagct	ctacagcgtg	gtggctctaa	98460
tcggaaattt	ctttttaatc	atagcaggag	tcaccaattag	cgtgttggtg	aatctttcaa	98520
gtagagtggg	agttccgtgg	ccacagagag	cagaggcaat	attcagcata	aagccctaga	98580
gaaagagggtg	ttgtgggcct	gtgcacacat	gtgtgtcaac	gcacatgtgg	cttgtggagg	98640
ctggcttccc	actctcaaga	tgaggtgtgt	gcaccccgag	ccttttgatt	ctcaaagctt	98700
tattaggacc	agagggactg	tgtgtgtgga	gggtgtgtgc	tcacagtgc	gaaacccaaa	98760
cctggcttct	ccaggagccc	acatgccaac	aaacaggctg	cacactcttg	ctagtacatc	98820
ccctaaagggt	atggggatga	gggaccaagt	gctttgcaag	acagcaggca	cagagtctctg	98880
ggacgtctct	gtaccccaga	ctcagccgcc	acccagggcc	agctctgatc	tggcttgacc	98940
tacttttctc	ttgtgttgtt	tttggaaagt	ctgatgtcaa	tgcagaattc	agcagagtgg	99000
ctgagtga	aaaaagagga	gagggaggaa	aagggggggg	ggacgggacg	ggccgaggcc	99060
aacaggaaag	ggcaggcaac	aagacaatga	ccacaaggctc	cctgtaacta	cactaactgc	99120
ttacctttcc	tgacccccag	ggcttagcca	atatagctga	gacccagtct	tgggtgctgtg	99180
gcttcagggt	aagtaaacag	ggaagagtgtg	gacatgggtc	tccattctct	ctcctcatcc	99240
aacaagggga	ggaggcagtg	gccaggcagc	catgccacc	gatgccatcc	ttctgggagg	99300
agccagacat	ttcaggcacc	tctccttccc	tgggtgccta	gaggtgctgt	gtctgcatcc	99360
atctgccatg	cctgccatct	gagagaggcc	actgggactt	ggtagagagg	ttctccacac	99420
atgctggcct	ggaggaaatt	ggtcttttagg	gacactgaag	gcagtttctt	ctgttcagtg	99480
gctccttgga	aacccacgtg	acagagctcg	catgacaact	tgccggctct	caactcccat	99540
tcttagctgc	ctcaagcact	gtaaggttta	ggagagcccc	agatgtaagt	atggatggga	99600
agaccctcca	gggagtcatt	gcctaccctt	ctgaactcta	acatgggtcca	gcttttccat	99660
tccacaattg	aggagacgcc	agacctggca	ggggagcaag	cctttgtttc	tgacccattt	99720
gcaaacccca	gccactgagg	aacttgcata	caagaaactg	cctctgggcc	tctcctggac	99780
tgagccctgc	ctcccagggg	acaactgggc	aacagatcct	tccagggtggc	tgcaagtga	99840
gatccatgct	tttatgacat	agaaaggcct	cagtctcagg	atttcacaca	ctgtatttcc	99900
ctcatcctgg	ggaccagggg	aggcgagcat	cttctgtctc	ccccaaacaa	gtgtgggaat	99960
gattaaaatc	attttttttt	tctgtcccat	gaactcatac	agttttcaga	taccgaggag	100020
acaaagccct	cctgtgctga	aattagacc	cgaaaaatag	gttagctgac	aattacttgt	100080
ttctaagtgg	agtgtgatgt	agtggcagga	gcgcaggatg	ggctgccagg	gctgcagtct	100140
cccccccccc	aaacttactg	tctcttaacc	tctcgagtcc	ctgggtttct	tgccgggatg	100200
ataattctcc	ccatctccct	cctctggtgg	gctgtggaa	agcgtaatga	atcaacgctt	100260
gaagcacgct	gaagaggcca	gactcgggat	gccaatgtaag	tacacagcat	cgccagccac	100320
ctctcaagtc	tacacggagc	tgatttattt	acctcccgtg	aaagagacaa	caatcatcat	100380
atttacactt	catgccgcag	cttcctgcgt	ggcacggcag	cacccccctc	ctctccgctg	100440
ctgaggactc	catcaagcac	gctgccttgc	caggatgaca	gcagcccact	ctcagcctct	100500
ccctggcctc	cttacagatc	atgacctctc	gcccggtgag	gtctgtcacc	cgaaaaccac	100560
ggtacacccg	gggctgcagc	ctctctatgg	gggaggctga	ggaaatgaat	tccgtaggta	100620
aaaggcttcc	taggaaatca	gacgtgcta	gtaattaagg	agcgaagcat	aggtgcgtga	100680
aaggtaaatg	gatgttattt	aaatgttgcg	tcatttaaaag	agtgtcctgg	tgcttcagtt	100740
ccttgttacc	atgcagggtc	gtggacgggt	ggcaattagg	ctggcacggg	tagagctcac	100800
ctgctgagct	gagggagggt	ggggacacac	cttcgggtaa	ttgctgctgg	gcagctctgg	100860
gtctcccccc	ccccgcccc	gccctcactc	cccccccccc	acttcttttc	tgacagctct	100920
ttcatttgca	gcagcttaca	gggcttggtg	cccttaccac	gaaaatcacg	ttggaagaaa	100980
tataagaaaa	agaggaatga	aagagaaaagc	cagaaaagtt	catattaggt	tccgatctgc	101040
ggccaaacct	ggccgagaga	atccatgacg	gtccgcgcgc	atataaccct	gtggcaacag	101100
ggccccggc	aacaggggcc	gccacaagag	cttcttgagt	tgccacctgc	caggagacag	101160
gatgaatgaa	tggatcatct	gtccttagag	cacaagccag	gcctgattct	ccaatattga	101220
tgtgtgaggg	agatgtcaac	agagggtccc	taaagaatga	tgcttctatt	tccatgctaa	101280
tcctggggcg	tcagcttcag	tcggaacagc	cggaccgtta	ccttagctct	gctgttctcc	101340
tgtctgtaac	ccgcagaggg	aagggcgggg	tcacccagca	ttgccactcc	ccccaccctc	101400
acgtggtcca	gacccctctt	gggttgatct	gctcctgaaa	aacagtgttg	gctcaagttt	101460
gcctctgaag	gtatgtcacc	gctggctcag	ccagcttata	tccccgggtg	tttcaagatc	101520
aaaaacacca	aacgaaagaa	aaactttgtt	tcaagagcag	agtgtgggtg	caactctgat	101580
caaagtgttt	ttcagcatga	caactcactg	cccgtagaaa	ccagtacttg	gctgttgtgg	101640
ctcagagtga	gatgcggagg	gaagtggatg	acaacagctg	tatccagggtc	caaacagagt	101700
agattcacgg	ctggcagaaa	atggctgaga	gccttgggct	gcattccctcc	tccccctctg	101760
cctctctctc	ttttcaagggt	gggttttgga	aatgtccttc	ctgtgggttg	tgtgcctttt	101820
ccatgttaga	cctggggcct	gtgcagatgg	ccctgtgttc	ctggtgctgc	tgttgagatg	101880
tgaacgagtg	ataggaaccc	aggcactaaa	cacacaatgt	ggttgtatct	gactagaagc	101940
aaggcaagag	caggaggcat	ttgagggtaa	aggagtgtaa	ggactgtgta	aagagatgag	102000
ggttctatct	gggaggcagg	agtcccaatg	ccagcaataa	caatggactc	tcctggctga	102060
cccaaccaga	gagaattcaa	gatggcagag	ggacaggctg	tctgagtttc	ctatggctgc	102120

accgataaaat	ggtcataaagc	agagtagaggg	aaaaccacag	acagaaattc	atgccattga	102180
gactagaaat	ctagctcaag	gttgtgtgtg	gcagggttgg	ttcctgggtg	ttcaaccttt	102240
tcacactgtg	acatgatgct	gtgggtctgca	gatgtgttgg	gctgcatcca	tagctaccct	102300
gggacacatt	catggaccgc	aggttacaca	tgctatttaa	aaactccaag	ggaagggtta	102360
gagaaatggc	ctggtagtta	agtatgcttg	ctgatcttcc	agaagacctg	agctctgttc	102420
ctagcagcca	tgttgggcag	cttacaacta	actatgactt	ctgagctcca	aagctctctt	102480
ctaatacata	catacatata	tacatacata	catacatata	tacatacata	cgtacacaca	102540
cacacacaca	cacacacaca	cacacacact	ttaaagaaaa	aaattctggg	ttggagaggt	102600
ggctcagcaa	ttaagagcac	tgactgctct	tacagaggtg	ctgagttcaa	ctctcaacca	102660
catggtggct	cacaaccatc	tgtaatggga	tctgatgccc	tcttctgggtg	tgcgctgaag	102720
acagatacaa	tgtactcata	tacattaaat	aaataaataa	gaaagaaaaa	agaagaaaaa	102780
gaaagaaaga	aagaaagtgt	aaacgaggaa	aattcctaata	taaaaaagaa	agaagaaaaa	102840
aaagaaagaa	agaaagaaag	aaggaaagga	attctgaggg	agaatctgcc	ccttttctta	102900
acttccagggg	ctataggcaa	cctgtggcct	ggggaagctg	tagacaacct	gtggcctggg	102960
gaagctgtag	acaacctgtg	gcctggggaa	gctgtagaca	acctgtggcc	tggggaagct	103020
gtagacaacc	tgtggcctgg	ggaagctgta	gacaacctgt	ggcctgtggc	agcatcatgt	103080
caacgcctca	cctctgtgct	ccaatttccct	gttctctaag	gacacatgcc	atcaaatgca	103140
taggacactc	tacatcaaga	tgatcttgct	tcaagatgtt	taacaaaatt	acatctgcaa	103200
agacctatct	ttacatgtga	ggctactcca	caggttctag	acatattttt	gaggagccac	103260
catccaactc	actatgtgac	agagtcactc	agagatttgt	gtccaggaca	gactggctgt	103320
atctgctctg	agagtccctt	gcctgcccgt	gggaactccc	cagtggctcct	taaggcccct	103380
gaggactttg	gatctgcaaa	gccacatctt	ccaaaacctt	tttctctttt	tggagagcta	103440
ctctaccctg	aaaccttttt	ctctgaggtg	gcttttagag	aggcaggtct	cagcagggca	103500
ctgtgcccac	aagaagtccc	ggggagaagg	gacccaaggg	ccagtgtctga	actatcgctg	103560
agactgagaa	cattgtgtct	cacctaaaaa	cgggtggctgc	aaggaccaag	caggctctat	103620
aaatgtctta	ctgcctttat	tcttttccct	ccgctccatc	ttactcctca	tttttgtttg	103680
tttgtgtgtt	tgtttgtttt	cttctgagat	gtagcccagg	ctggccttca	gctcactatg	103740
taactaagga	tgactttaaa	cttctgatcc	tttcttccct	ccaactccag	agtccctggg	103800
caggtgtgtg	ccaccgtacc	ccagctttat	ttgagactat	gattcaggct	ccatacttca	103860
tgcatattag	gtaagcatgc	taccaacttg	gctatattcc	cagcctttct	ttctttcttc	103920
tttgagacaa	tgtctttttt	tttttaatta	tatgagtaca	ctgtatctgt	tttcagacac	103980
accagaagaa	ggcattggat	cctatttaga	atggttgtga	gccaccatgt	ggttgttggg	104040
atttgaactc	aggacctctg	gaagagcagt	gatgtctttt	aaccgctgag	ccatctcgcc	104100
agtccttgag	acaatgtctt	gctatatggc	acatattggc	ctcaaactca	gaatccttcc	104160
gcttcagcct	cctaaatact	gggattacat	gtgagccatg	gtgtttggct	tctagccttt	104220
cttccctccc	tttcccttcc	cttttccctt	ccctttccct	tttcccttcc	ctttccctcc	104280
cttcccttcc	cttcccttcc	cttcccttcc	cctcccttcc	cttcccttcc	cttcccttcc	104340
cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	104400
taatcctggc	tgtcttggaa	cttgcctctg	agaccaggct	ggcttgnnnn	nnnnnnnnnn	104460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	104520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaaga	agaagaagaa	gaagaagaag	acaacaacga	104580
cgacaacacc	ggcgccgctg	cctccactgc	catccacctg	agacaggact	caaattccaga	104640
ccaatttttt	aaaccagtg	ttcaagccgg	tagtagtccc	tagtagtccc	acttgggatt	104700
atagctcctt	actttgtttt	gctttgacgt	tttgtgacat	ggtgtgatgt	agtcttggct	104760
gtcctagaac	tcaatgtgta	aattaggctg	gccttgaact	tgcttctgcc	tcctgtggg	104820
atgatagact	gatggtgtaa	aactccactt	aggaggcaga	ggtgggagga	tcagaaattc	104880
aaagtcatcc	ttggctatgt	tgtgagtttg	aggaccaacc	ttggctacat	gatattcctat	104940
ctcaaaaaga	aataaatgta	ttgcggggca	tggtggcaca	cgcctttaat	ccagcactt	105000
gggaggcaga	ggcaggcaga	ttttctgagt	tcgaaggcag	cctggtctac	agagtgaagt	105060
ccaggacagc	cagggtctaca	cagagaaacc	ctgtctccaa	aaacaaaaaa	caaacaaaaa	105120
agtgaacccc	aacagtactg	ccggacagtc	tggtgtcttt	cctaagtctc	ctttcaactc	105180
tgtttaccca	ggtgtaccca	caagggtgtg	gagcagctct	ataccagag	gtgatacggt	105240
tgtttgaatg	agagaaaagt	ttcccatcag	ctcgggtgtg	tgaatactcg	gtccccagtt	105300
ggcagtattg	gctggagagg	tgatggggag	gtgtagcctt	ccgggtggaga	tgggcttttg	105360
gagtttaaa	cttccctcac	gaagaccact	gggtgatctt	tgttaccaac	agaattggat	105420
tggcctgctc	ttggctccgg	cctcagttta	tctaaaattt	acatgttacc	tgatcaaaaa	105480
ctgtttcctc	ccccacccct	ctccctgtct	gtattccctg	ccctctagt	gtgctggctg	105540
tatactacac	tggtgatctt	gactgtattt	cagtttacct	cttgttctct	ctctgctgac	105600
tctagagatg	tcctttcatg	gctgggacct	ggctcagaaa	tttctaaggc	actgagcctt	105660
cctccatctg	aacttttaga	aacttcttgg	cttaaagggtg	tatttctgac	ttagtatgca	105720
atgagacctt	ggagcctgca	ctttgttaag	caccctgggt	ggtgggtggg	gtgggtgggtg	105780
tggtgggtgg	agtagtagaa	ctctgatgaa	cagttagtta	ttcaaggccc	atctagaaaa	105840
agaaaggctt	tgggtgcaca	tggctataac	tcagtggcga	gagacgtgct	tcccatgaat	105900
aataatgatg	acgatgaaaa	taattctctg	tgactgttct	ccccacttcc	ctctctctca	105960

ccctagctct	tatctaccga	atccctgcac	aagcaccat	ggggtttaca	gaatctgggg	106020
cggaacgtta	gtcacttccc	ttcgccact	tcagtattgt	gtttccagaa	gtaccattt	106080
tggctagtc	ctgaggaaaa	cggcagctgc	ctgtgggcca	ccagcccatg	ccaagtggag	106140
tcagcaagaa	agaagctgac	agcaaatgtg	ccaactgtgg	gtctgctgga	tttctactgt	106200
gctaagtgg	ttcaagaagt	ttcttcttaa	ccccatacaa	gaaaccacaa	atthttattat	106260
ctacactggt	ttgtagatga	agaaaacacc	attccgaagc	tcactgccag	tcagcactgg	106320
aactgggaatt	tgthttggcta	attcagtggt	tctcaaccag	ggtggatttg	gactccccag	106380
gggatattht	gggacacttc	tggtcgtcat	aattgggac	atgtgctact	ggcacctagg	106440
gtagaggcca	gggtgggtact	gaccttccta	ctatgaacag	ggcagcca	tcacataaatt	106500
ctctcgatca	aaacatcaac	agtgcctaag	ttgagaaatc	tcaggctgaa	accctgtcat	106560
ttggccttga	ggtgggtggg	aggaggttag	agagtggaa	aaaatcagaa	ggggccaccac	106620
ataggcctcg	agtggaggag	gaacagggtc	cctatgctag	ggataatgga	gaatagggca	106680
gcttggtgaa	actthttcttt	cttccaagct	tggctagagc	cctgctcaat	ttcccccaac	106740
tctgccaagt	cagtccccgg	acttcgcact	aagthttgtcc	tggagtgacc	ctgactccag	106800
cttgagagct	gggcaacaca	tttacctggg	tcttcccagg	agtgggttaa	aagtcaaaaga	106860
taagtggcat	ctgagaagtt	aaaagtgggg	tgaagcagca	aagagctggt	gaccccatca	106920
atcagctgac	ctaggaagga	agagagcaac	agcctthtct	gtthtaggctt	ccaggagact	106980
ggattctgac	ccactaggct	tatcttccac	agcctthtct	gtthtaggctt	gggctcagtt	107040
tcctcgcac	atccgcagga	gcccctggag	caccacatt	cagccggccg	ccaggacagg	107100
ctccccagca	gtggcctccc	cactaactga	cagtggtgac	aggaaatc	tcccatcca	107160
atctcctcag	aagtctgaat	aagtaaggga	cagatgttgg	ggagaggcgt	cactctgggg	107220
ttgatgaaga	aaagatcatg	agaagcatat	atthttaccg	ctatggttgg	ggttccatta	107280
ccaccctatg	cggggttgag	ggggaagggc	agaaaaaagg	agatggagaa	ggacagacac	107340
gtaagaagga	ggatgtggtt	agggtgatt	cgtcccctgg	ggtcgaacaa	tcagctgtta	107400
ctggggcgga	aggcaggaag	tctccttaaa	gacacaatat	tctgaacgtt	gaactcagga	107460
tttgaagcaa	gcccagcagt	caccttagtg	gagcccatat	ttaatthaac	acagaagcgg	107520
ctatctcagg	cttccctctc	atthcttht	ctcagatgct	ctcacttaga	aagcttagat	107580
gctcttagaa	atgactcaaa	agtcaagaa	cccaggccac	aagthtctct	ttgggggtgg	107640
ggagggtg	aagaggtggt	cccagctctg	tccctthtaa	taagcaattc	agcagcttht	107700
gccaagtcat	tgggttcatt	tgggttht	ccccctccc	gcctthcaga	ctctgattgg	107760
cccctaggga	aggagccg	tcttcattgg	tctccacctt	tgaaatcact	tccctaagta	107820
ggcctgagtc	agagaagcgt	ttcggagggc	gggactgaat	gggtgttaat	cttagaacgg	107880
ggtthtctg	tgatactact	ttggtaaaga	tcttccccta	atthttaaaa	agacgtctcc	107940
tctctaaaa	tgaggcgaa	tcttht	agaacgtgcc	ccttgagaag	ccgtgggctc	108000
ttcagcgact	aagacgagac	attcactaga	aaagatttca	ctaaacccac	gagggataga	108060
ctagacctcc	agtgaagatt	gggctgtg	gggtgacatt	tgctccctata	ccccgaagac	108120
ctcgagctag	ctctccagtg	aagactggg	ccgtgcgagt	gacagtggte	ccctataccc	108180
gaaaaaaaa	aagtcctatt	tgtggaaaaa	aaaaaagact	tgggtgtt	tgctgcatcg	108240
gtggctggct	tccatcttta	gttctactca	ctcctgttgc	ttcgctgct	ccaccttcgc	108300
ttagctcagg	cctcctgtga	atcagtht	aggctaaaa	aagthtccaa	aaggaggggc	108360
tgtagccctt	taaggacttc	cccgcgaccg	agtcagagat	cagthtaaaa	atgccaactc	108420
acagagcg	ctgcattctg	ggaagctgag	tgtcaccgta	agaacttcat	tgaccggaat	108480
gcaactgcaa	aatacacgcc	tatacttctt	tctgctctt	aaactgtagt	ttgacgtaaa	108540
gctggtctaa	gcaagtcgcc	taggcccagg	gttagccaca	cctthtcagc	cattggccag	108600
ttggttagtt	ggtaggcg	gcttagagaa	gctcctccag	gcaaggggg	ggcctcctt	108660
ccaatcagag	cccagacgcc	tgaatggcg	ggagtaagca	gaggtgctgg	cgcccccgag	108720
tgggtgtggt	caggttgccc	agcaatgggc	ggtgattggc	cctgggtggt	tcattcgag	108780
ctcgtgcgtc	acgacgcgc	cagctgatcg	gagactggag	ccgtgtgtgt	ctgggcgtg	108840
ggaagagaca	gagcggtcg	cagtcggac	aggctgcagt	gattthtgc	ctctgtccac	108900
agcaaccccc	gcacccagca	tcaggtgggt	gtgatctggg	gacccggtca	tcccgggggg	108960
aaccgcggt	accgggtgat	ggggaagta	gggtcctgac	ggccacaccc	tgcccttctg	109020
ggggagggga	gagggggcg	cggggacagg	ggcgtctt	ggagaggagc	ctggactctc	109080
ccgagtagtg	tgtctggacg	tttaaagaga	gagtcaccga	caggagtcgt	ggcagaaggt	109140
ttggagaagt	aactggggag	gaatatgaga	ggccagagg	ccgggggct	ctaaccgccga	109200
cgcccttht	tttgaggatg	cccagctga	ccatthtagc	tagggaggat	ctggacgagc	109260
gaggggtg	gaggtgcatt	gcctctaccg	gcgtgactg	ggtcagggcc	agttcaagtc	109320
cctggcaggg	aaggggtcgc	tgggcgggtc	ggccctctct	ctcgttccct	cccggggatg	109380
ttatgtaagg	ggggagggga	aaggagtagg	ggcgggcggt	gcggaggcct	tatgcaaccc	109440
aaaggttagg	gtttcacccg	gggttggcg	gaggttgggg	ggggcgga	ggaggagtgc	109500
ctggaaactc	taccgcaccc	ccccctccca	gcctaaactg	ctgtcttggg	cagagagaag	109560
gtcaccttht	cacctcccc	ctagtatgtc	cggtagagag	gcccctagcc	cggtcttggc	109620
ctgactgcct	gggaagccgg	ctggctgggt	ggggcgccct	ggttagtcat	cgctgggctc	109680
cctctctccc	cacctctcgg	ccaactcttg	gcccctcccc	acggcctccg	gttaggctaa	109740
cgttcccacc	tccctctggc	cctagthtca	gtctcccaact	catttggcct	gtcaccttgg	109800

ctgttagagt	aggctagaag	ctgtcatggt	gccagagagt	tgatggagca	gctgggtcaga	109860
gggtcagtg	cctgggcccc	ccccgcccc	cagccaagg	cacctgcttg	gcacaaactc	109920
tcagcagcca	gtgaaccttg	tggcctgaac	agagctatcc	tgggcagaga	gaagtggaca	109980
gagactgac	acctaggaga	aggaagatcc	gacaaagttt	atacttccca	agaggctttt	110040
ggaatttgaa	agttgcccac	cctagtgtaa	tctttccact	ctctgaaaat	agaaatccca	110100
aggcaaagtc	tccttgggccc	ttctatctgg	cagtggccat	gtccttggac	tgactgtgca	110160
gaaccaccct	ctcgggctcc	cagccctcta	gectgccacg	ccccagccc	cctccctgag	110220
ccatgctgta	gggccccggc	ttttactgct	gattcatgcg	ttggaactgt	ggggggcggg	110280
cttggaactt	ggaacaaagt	tcagacgtgg	agggggccg	agacagcctg	gaattcatatc	110340
cagatgtacc	cggaatgtgc	aagcggaatg	cctggcatct	ctagtctga	ggaagctgcc	110400
cagccaccct	accatacct	ccctccctc	ctggccttgg	tcagctgtcc	tcctcagac	110460
tcctgagagc	ccctgctgac	cttccaaactc	tagtgccct	cccatttcta	accctacaca	110520
aacctcctt	gctgctgaat	tccttaagaa	caagtcattt	gagttgatca	cagagctcat	110580
atttctgaag	tacatttttt	tttttaactt	gggacttggg	ttctacaccc	tgccctttga	110640
atgccgaaga	tgctgggctc	cttagcaggt	tgccaagagt	tgccagctcc	tagtctgtaa	110700
aggggcacaa	agcaagtgc	tttagaagcc	ctttgcttct	tattcaagaa	ccctcatta	110760
gaaggctactg	aaagtacgt	agagccaggt	ttggatggcc	tctgggtcgc	tgccctgtc	110820
accagcttt	cctgtttttt	tttttccctc	ccttcccttt	aggaacctgt	gcctcccaca	110880
ccctcacctg	gctgagccgc	agtagttctt	cagtggcaag	ctttatgtcc	tgaccagct	110940
aaagctgcc	gttgaagaac	tggtgcccctc	tgccccggc	ttcgtggagg	aagaggagaa	111000
gcagcagctt	tgccctatcat	ccggaaggtg	acagaactgg	ggtgggaagg	tctggacagc	111060
tgggggtgat	gctttatggg	agggaaaccc	tggtcctctg	gggagccctt	acccccactg	111120
gcccagtgaa	agatttaggt	taaaggcact	gtctataaat	tggggaatag	gtgactccac	111180
ctccccaaaga	ttagttgatg	tctgtgtggc	agtgggaaga	aatagaagga	aaagtctgtc	111240
tgtttactga	gacttccctg	taggcctgcc	tttcttatct	tcacatcac	catgccaaca	111300
cacacacaca	cacacacaca	cacacacaca	catacacaca	catacacaca	cacacacaca	111360
cacacacttt	cctttccatg	aggtccaaaa	gtaaatgtac	tcagggaagg	ggacattgaa	111420
actcogttct	aagttagtcat	ttgtgtattt	actttttttg	tttatttgtt	tgattgactt	111480
tcgagacagg	gtttctctgt	atagccctgg	ctgtcctgga	actcactttg	tagactaggc	111540
tggcctcgaa	ctcagaaatc	tgccctgctc	tgccctccaa	gtgctgggat	taaaggcgtg	111600
tgccaccacc	gcccggctgt	atttacattt	ctttatttat	ttttagtctg	gcccagattt	111660
tgggtttagg	ggtacttacc	cttacacctg	tggatttata	cacctgtata	atggggaatc	111720
ccatagataa	gtaggcagga	gggcattaaa	agtcaccag	tggtgactca	gagcctgggc	111780
tcttcttctt	ctcgtggatg	gaaacgaaac	agctcttcac	atgaactgtt	gtccttcccc	111840
cacccccctga	ctactcacc	agctcagggg	gattaggatg	gaaggaaagg	ctatgggttaa	111900
gtcccaggca	agctcgtggg	aggctagtcc	tctactggct	tctcaccatg	catgggtggt	111960
ccaaggcttt	ccctccacct	aaagcaaaac	tgtagctctt	ggttgggttc	tagcaaccac	112020
tgccatttat	tttctgcctt	tgctttccag	gatagtga	ctctgctcaa	tactgtgcag	112080
gcaagaaatt	gtcaggggag	atgggttgta	tgatatgagt	cccttctgct	gcctctagct	112140
cctgattcat	tctcacgtat	gggcttggtc	tctgattgtg	gttcaccttt	ggcccagctt	112200
tcctaacaga	agatgggttc	aggggttaca	ggaggctgtt	tggtgtattt	gacaggagga	112260
ggagttctag	cctgttcccc	atttgtgaga	aactgaaagt	cataggggag	actagatcat	112320
ctaattccagc	cccactgcag	tctaagctga	gggataggat	gtgtaaggga	ctgtagcaga	112380
cgggctgggg	aggctgagtc	ggctcacaca	ttgcgacaaa	gattgccctt	ccctcgacct	112440
cgcttgcttt	ctttcctcct	cccttccctg	gccacagtgt	gtccctccag	cactgggtac	112500
atggctctgc	tgctcctcatc	caacatggag	cctcagaggt	gagaaagggc	agcctggaag	112560
caacagaggc	aggcacaaga	cagtggagga	cctggcctgg	aaccacaagg	gcctatccgg	112620
acattggtca	gagaggcacg	tagaagcctg	gagaacacca	ggaaagagag	cagccagcca	112680
gcctcagtga	aagacacgtg	cttccagcca	tctcctctca	ggacctgcct	tcctgggaga	112740
tgaaggccct	ccaggaagta	tggtcccatc	tctaccctgc	agtttctata	aacagcctca	112800
aggagcatga	gccacctctg	aaaggaaata	cacagcaaat	tcaaaaagag	attcaaatgt	112860
gtaacactgt	gggaaaacat	atctatgact	ggggttgtag	ctcagttggt	aggtttgctt	112920
aacatgcacc	aagccctggg	tctgtcttct	gcattgcata	aaactgaaca	ggttggccca	112980
gtgctataca	atgagttctg	agccagcctg	ggctatatga	cctacattca	aggtaatcct	113040
caaaaataaaa	taaagcattg	gttagtaatt	caaagaaagc	gactgtctca	aaaaataaaa	113100
tcctgatca	taatacaaca	agcaaatgaa	agccagaaga	agatgtggct	gaaaccgttt	113160
tggcagttacc	aaccattgtg	agagatgcct	ttggacctcg	aggctcctgt	gccttgtgtg	113220
gtatcctaaa	ataaggattt	ggttataaaa	tgttcatctc	tagtttgcctg	tcttagaaat	113280
tggaacgcag	ctgtttgttt	ggaagtccat	tccttttctg	agggttgtaa	tagagaaaaa	113340
tagggcttgc	ctaagtaaat	tatatccatc	tgatggtggt	ctgtcatgaa	aatgtatagc	113400
gtcttacaga	agaaaaattg	agtggaaata	taaattattga	gttctgtgca	gccatccaaa	113460
gtatgagttt	gtgactgttt	ttaaaatatg	aacacatact	atactaaaaa	gattataaaa	113520
accatccaat	tgagtggaaa	tataaatact	gaatactaaa	tgtaatatat	tttttataaa	113580
				aaagtatgag		113640

tttgtgactg	tttttaaaat	atgaatgcat	acttgtaata	tatttttttaa	aaaaacactg	113700
aaagtggatt	caaaatgtta	agaatgggtg	tttttgtatg	gtgggatagt	acaattgtgta	113760
attttccctt	tgttttttct	gtctttctaa	tttttaaaata	ttgtgcattg	ctttcatatg	113820
ttaaaataaaa	tacaaaagac	aaataaatgt	tttaaaattt	ttactctttt	atgagtgttt	113880
tgctgtgag	tggcaggaat	ccaacattgt	cttctgaaag	cagctagcgc	taacttctga	113940
gccgtctctt	cactccctct	gtaattttta	aaaaatatat	ttgtatgtta	tattatgtgt	114000
ctttgtgcac	cagagtgtgg	gtgcacattc	tgcagaggcc	agaagagggc	atcagattcc	114060
ctggagctgc	acgctgtttg	gatcttctga	tgtggatgct	cagaatcgca	ctcaggtcct	114120
ctagaagagc	agcaaatgct	cctagccact	aaagccatct	ctccctctag	tcctcattgt	114180
catgtttaga	ttttggagaa	tttgcttgta	ggaggatggg	ctacaaccaag	tgccagggtga	114240
aagaaaatgt	ttgcttgagg	tacctattgc	ttcttgagt	tgtgtgtgca	tgcttgtgtg	114300
tgtgtgtgta	tgtgtgtgtg	tacactggag	ctaggaatca	tatccagggg	ccttttcaag	114360
ctccaccaca	ctaaagtcaa	ttctatgaac	ttcattaatt	gtctgaatcc	acctactctc	114420
tacacacagg	aagcattcct	ctgactttct	gactgtcagc	cagctaagga	gggtgtggctt	114480
agaataagaa	agaagggaaa	tgctcaaaaac	ctgtcactct	ntggggnnnnn	nnnnnnnnnnn	114540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115320
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	115980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116340
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116400
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116640
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116700
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	116940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117060
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	117480

gacccctaac	ttccatctct	tgggtcccat	cttaccatt	cccagactcc	cattctgtgg	117540
ctggggcttg	ctggctgggtg	gtgtcgctgc	agaggctggg	gaagggactg	cttgctgctg	117600
ctgctgctgc	tgctgctgct	gctgctgccc	atatcctgga	gatcaaaaca	ggccaattca	117660
gtgcaaaccc	agtcaagaga	tttcaacacc	agcagtaaat	accttaggaa	aacccacctt	117720
tggctcttag	aggagcagac	tgcaccaagc	tgcacacctt	gtgtctgcaa	agggtcact	117780
tttttgtctg	cgcatagtaa	agactgggtc	catttccctt	tcaactgttt	agaattgaca	117840
gctttcagtt	ttaatgaggt	tctccctgaa	gccttgccca	tttctccttc	aagaacctgc	117900
agtaaaagtc	ctgggtctca	tctgtgatac	taccactatt	tccatagcaa	caaatcgatg	117960
tcatcaacag	gacgggtgtgt	tctgtggcagg	atccattaga	agagaaagca	gggtggtaag	118020
gaaaacctaa	aagcagttca	attgtctctc	aagtgccttg	gcttcaaagg	aaaaggagac	118080
gtttacaaag	gggtctaccc	tttcccccaa	acaaagcaac	cttattttgc	aactgacagc	118140
tgggtgaaaa	tatgagttga	aagagtctca	ggtagcgtga	agtgtagtgt	aactgtcaca	118200
gcaggagagag	atgagaacag	agttcagaag	cagccgtgtc	aaggcaacag	ggagataaaa	118260
aggaggggga	ccgagcacgc	ttctaaagca	gcaaataaga	cgcgccaccc	tgctggagtgt	118320
tgtatttgcc	cttcaccttt	tatacaaatg	tttaccagt	tgaagaatgt	taacattgta	118380
aattgcctac	cactattttct	aaaataacga	gtttatgggg	tttgatttta	cttctgtgat	118440
tggctctgaa	ggctgaagcc	agcctcgttt	aggctgctca	cctggagtgc	ttgtagtatga	118500
caaaggcagc	taaaaaaaaa	aaaatgtccc	agagctcctg	aaacactaaa	actggtgcgc	118560
acaggcagga	agtctccctt	gccactgagg	ctctcccttc	tccaaactgt	agttctaact	118620
cctgtctggc	ctcctgaggt	cccaactcac	cgggcatgct	acatgcccat	ggaattcatt	118680
tttgcaaaaa	ttactcttaa	gtagctcaag	gaatgagacg	taattgtgtt	gctggagaca	118740
aatacattca	gcacaaagtt	gagaagatta	aatagaattg	attatgcttt	agttcatcct	118800
acagagagaa	aaagttagac	acctaatttac	atagagagag	ggcccagggt	actgacaatg	118860
agcccttccct	ttacagctca	gccttcccat	tatcactcaa	ctcagacacc	cagcccagtt	118920
cagtctatgt	gaagcatttt	taaaatcagc	aagagaaagg	tgagtgtgtc	agtagtacac	118980
tgaaggaaa	tagaaatgag	cacacagtga	cctgctgcat	aagacacagt	ttaaaagggtg	119040
acctatcttc	cagcgaagtt	cttgcccttct	ttaaaaagaa	tgtgtgtattt	gctggtatgt	119100
gcacctcacg	aaggtgcttg	cagctatgtg	gagggcagaq	accatcctga	gggaatcggg	119160
tctccttcca	ccagagtcct	agggatggca	caaaggcacc	gggcttgaca	gatgccttta	119220
cctgtctgagc	cctttcccca	gcctctcctt	ttctacagtc	tccaaattac	ctgaggtagg	119280
agcctcattc	tcagaacccat	ttctgtctgcg	catgtctgac	tgaactggat	atggtcactc	119340
gactcaattt	ctataaaaa	ataactgagg	agccggcttg	gttggtagag	aaaacataaa	119400
gacccgagtt	ttgttcccac	atggggcggtg	cgtgtagag	aggtgaggag	ggactaggca	119460
gggagcagag	gtaggcagat	cccaggcttc	aatagccagg	cagcccaacc	tgagctctaa	119520
gttcagtaag	ggagcctgta	accacagaag	acactgtgcc	gtgtaaccac	ggcaagcggc	119580
ccgctgcagt	ggcatctgct	catagccaca	gctaccagag	aggctgagcc	acaagactca	119640
ttgaagctgg	aggtcaaggc	cagcataggt	accataggta	gacccccatc	tcaaagttga	119700
acagtaaaaca	tatatatttac	tacaattaaa	aaaacaaggc	cggggtgtgtg	cgcgcatacc	119760
tttaattcca	gcactcggga	ggcagaggca	ggcgatttct	tgagttcgag	gccagcctgg	119820
tctacaaagt	gagttccagg	acagccaggg	ctacacagag	aaaccctgtc	tcgaaaaacc	119880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaccaa	aaacccaaca	catgtacctt	ttctaacata	119940
aaattttcat	tttttataaa	aattacagtt	ataattttta	gttgaacata	atacaatgat	120000
aagtctccaa	cttgataatc	tgaggctggg	ggtgtagctc	atttagtaga	gtacctgcct	120060
agcagtgcgt	aagtccgggt	agtcctctgc	ctgtatccca	gtgcttgggg	aacaggcaga	120120
aagaggacca	gaagttcaag	gtgtcctctc	tcttcaggta	atgaggagtc	tgagccagcc	120180
tgggatgcgt	gagacacacc	agtaacagca	actactgcct	ggagtgtctc	ctctggacca	120240
ggaacaagag	ttaagtgcgt	tactcccgtg	actgtctgca	caatgatgga	gacagtacgt	120300
catctttaac	atctttggct	gagaagagaa	aacctgggat	tcctcagctc	gttctggcta	120360
agttcatgta	cttcatcatc	atgcagttaa	tttgtaaaaa	accaatcctg	gcagcaataa	120420
ctctaattat	atagaacata	agatgtggta	attaggaaaa	gctactaatc	cacttaatat	120480
agtaaccttt	atctcttgca	aatctgtgtac	aagagacagt	cccaaataca	atgatggcaa	120540
gaattccaga	gcattgtaaa	tagcagcaat	tgcccttcaa	ttaacgcatt	gtaacgcagc	120600
agctgcccac	aagacctcaa	atcaatcagt	ctatagctaa	ggaaaaatct	ttctaagacc	120660
aaaaccattc	tacaaagcag	taacgtaggc	tccgtttata	ataacctgtt	ttgggccacc	120720
tgcaaatcaa	gctatcccag	gaagccagat	cgtaattctt	aggctctgct	ggctacacac	120780
tgggtcccaag	ccatgagggg	actagattac	agcaggctcc	gccctcgggt	acctgctcat	120840
agctatcatt	cttacagctc	attatggcaa	gtgagctctg	ggcagagaaa	aattcacaaa	120900
caaaccaccc	aacttcccaa	gcaagcattt	tcttaacaaa	cacaaagaat	aaataaatag	120960
agcctgccat	ggtagtgcac	acctgtaate	tcagcatgtg	ggaggcagag	gcaggcagat	121020
ctctgccagg	agttccatgc	cagcctggtc	tagacagttg	caagatcaac	aacgctatat	121080
ggtgagggcc	tgtctcaact	cccaaaccac	tgaanaacaag	taagacgata	tggatcaatg	121140
caatattttct	cagttcctac	tgacagaaaa	tggacacaat	taggctgggt	ataattccaa	121200
tatgaataat	aagtatatta	tacagtacct	agtttaagtc	tgagcaagat	attatcgcca	121260
acagcacaga	tacaggacac	acacacagct	gcaagcgtga	aggatttaaa	ggcaccctac	121320

cctacagtat	atgcagcatt	gactgtctag	ttttattttcg	cacactttga	atcatccatc	121380
cattttttgct	caatatggca	gcagtaataa	aatgtatatt	tgcattttga	tgcatgggtg	121440
aattcttact	agcctgggct	gtgtttgctc	actaactcca	gtggacttct	gacgtaagag	121500
gcgctggact	agcttccaga	ggatgtaaat	ctaacttttg	ttctcggcct	cccctgaagc	121560
ctttgctgtg	gtgaaagggt	ctgtttctga	agccacgaca	gtcccatggt	gggttgagta	121620
acaatactcc	tggcttggtg	acaatgccaa	aaaaataccaa	aaaaacccaa	acccaaacca	121680
aaaaaagcac	agagctcaca	tctgagccaa	aaaaaccgac	actccctatt	ttttgaagaa	121740
ctcacagaaa	tcaagaagaa	aaacaagcaa	acaaacagca	cataacaggc	taacaacaac	121800
aacaagtcca	cttcagaggc	caaaaaacca	aaggcgaagg	ggctcttaaa	catttacaag	121860
gacagacctc	actcagaaca	caagctatag	ccatgatact	gctcttcatc	catcaattct	121920
taaagacaca	gaagatgggc	tgctggcatg	actgggaaga	gaccagtgtg	ctcactcatt	121980
gctgggggtac	gctgtaacaa	gcagaggaga	atgtctaaact	gtgtctgcc	atcacaggca	122040
cattcccagt	taaccagca	atcacttctg	gcaaaaaaag	cccactctgc	ccgcacactg	122100
gtgtacccag	gaggtaatc	actgcggtt	ctgggtggatc	ttttttcctt	cttgacgtgc	122160
ttgggttcaa	attaggatca	agcacacttt	tatagtcaga	gcacggagac	aaaaggatct	122220
aaccacagag	ttggttaaat	aaacaacaga	acagccacac	cagaactgct	ggggggcggtg	122280
gggagggggg	ggaggggaagg	gggagaagct	agaatcccag	cattcaggat	gcagagggtg	122340
gtaaacaaagt	ttcaggctca	gtcgggggtat	aaggtaagac	tttatttcac	aaaaataaat	122400
aatttttttta	aagaggacag	aaaaaagaca	gcgtagagaa	ctagctatct	tttatgtaaa	122460
ccatgtgggg	tacaggagca	gcatttgc	ttggttactt	ggcgaatagc	cctggtagga	122520
taaagaaacc	ttaaactgtg	ttatctataa	agggcagcag	tggtgtgcaca	atgcagtggg	122580
taggagccca	cccatgtcac	gccatagttt	tgattacaaa	gtcacgcagc	tctactcaaa	122640
aattaaaaca	aacattaatg	cttggttaag	aaacagggtc	gcagagatga	tggtcagtg	122700
gttaagagca	cagcgcccat	atggaggctc	tcagccatct	gcttctccaa	ttccagggtg	122760
agctaacgcc	ctcttttggc	cttcaagagc	actgcatgca	catggtagcg	ttacctacat	122820
gcccgcacaa	attcaaagaa	aaatacaaac	tgtctataaaa	cccatatatg	accatcttaa	122880
gattcttttca	tttttttgag	acagggtttc	tctgtgtagc	cctggctgtt	ccagaacttg	122940
ctctgcagag	caggctggcc	tcacaaccag	agatctgcct	gcctctgcct	ccacagtgtc	123000
ggaattaaag	gtatttaaca	cacacattat	acatatcttc	cttctcttct	ttcttttttt	123060
aaagcgtttt	gttggttttt	gttaagtttc	aattaaaaaa	ctacatagtt	ttataggcaa	123120
acataatttaa	aaatgccaat	gtgaaataaa	taatatatac	atatataaca	ttctgtataa	123180
gattcactca	cacaacttat	atacttaaat	acaattttca	caataatgaa	aagctttgaa	123240
atgaagactt	ctggatacat	tagaaaacgt	ccctgaaaat	cgcaaatgac	gggttttcatt	123300
tctttgtgtc	agacattagt	gtgagtgtct	aaacttgc	aaaggctctc	ttctctatca	123360
cttctacact	attgcagtgg	ttctcgctaa	acctggaacc	tggtggctcct	gttccttggc	123420
tggtactaaa	ggcagcaagt	cccagcaatc	ctcctgtctc	acctttcttg	gaaccaatgt	123480
tataagtgtg	tgtgggcact	agccttggtg	catggctgtc	gggactagaa	ctctggctct	123540
caatattagg	ctcaagagc	tcttaactgc	taagccatct	ttctaccctg	attagaattt	123600
cttgaagcaa	aagaaactca	cagatggtca	gagtttacac	acacacacac	acacacacac	123660
acatacacac	acacacacac	tcacacacca	aggcttagtg	accactgtga	aaagggaagt	123720
gcggtgagga	actgtaaaaa	taagggtgtc	ggaaagctct	gcacaaaatg	gtgtcctctg	123780
gacaagccag	ggcctctgcc	ctcatcagct	cttagtaact	atggttgctc	acagcaaac	123840
atgccaggga	ccactctaac	atggagctgg	gatggctctg	agaggccctg	ttattaaagg	123900
agaagctgta	gagagttgat	ttgatggatt	ctaaagtagg	gggaatcgat	ttcctttatc	123960
gatattgggtc	agccatgctc	tagtgagtgg	ccccacaccc	acctatgagt	atgtgtggac	124020
agcacacacc	ggacctggca	agtcaataaa	acaaaaacaa	aaacaaacaa	ataaacgttc	124080
tggtccacca	tagtggctgc	tcgtgggttg	gagctgtccc	gcacttatgg	caggaagaca	124140
gtggctgcaa	agaggacaaa	agtctctgga	actgatcaac	tctagagtct	gcttgttatg	124200
agaactggga	agtaccgct	gggacagaa	cagactctga	aggtgatcag	gacagagatc	124260
acagaaggag	agactgggtg	tcggaggaaa	tctgaaacat	aactcgacgc	atactggtcc	124320
aaactgggtg	ccatcactac	aacagcagta	attgaattgg	gcacaacatt	cagaaaacag	124380
aaaaagacta	cagagtacgc	acctgggcta	tcatacaacc	aggcgattct	ggcactattg	124440
gaagcaagcc	agactggaga	aaaggaaaca	aaaagttatt	taacaaaact	tcccagagca	124500
tggttaaaaa	aaaaaaaaga	aagccagaca	tggtgggtgca	cgtttttaac	cctagcactc	124560
aggaggcaga	ggcagggtgag	gcaggaggat	ccatgagttc	gaggccagcc	tggtctgtac	124620
agtgggttcc	aggaaagcca	ggcaacaaag	aaacctgtc	tcaaaaatca	ctgactgggg	124680
agagagggaag	tggatccaag	agcaagagag	agagagcaga	gagagtgggg	gtttggatgt	124740
cagcggttatt	aaatgacagc	agaaaagatg	ggcccagcca	atgacatccc	agaaatggca	124800
aagatgaaaa	aataaacaca	agttctaaat	atcattttta	taatggctgt	gtgtctgctg	124860
gctcattctt	gttatccaaa	caaaactaaag	caggggtggc	cggtatttac	ggccagccag	124920
gactagagtg	agacactgct	ttaaaaaagc	aatagatgca	cgctaaccat	taatcagcgt	124980
aactggagtt	tgagggaggg	agtgggccc	ggaagcctgt	gcctataaac	ccaacctgct	125040
aggcctgaag	cccgaaggtt	gaactgagaa	catctcaaga	caaagcacag	gcacaatctc	125100
ttacaaacag	tttaaacaca	ataccagcaa	taaattgtca	gctttatgac	agataggctg	125160

acaggcatac	cacaaagatc	gggaagaaga	aacgggattc	ttgcacaaca	ttttacaaat	125220
cgacacagct	gagcctagt	acaggccgtg	ataagctcaa	aacatgcaca	ctgcaaacac	125280
cacagcatgg	ccagagtgc	aggggtcacag	caatacagca	acagagacag	taagaaacat	125340
tggaaaaggg	aagaaggcaa	agagcaccac	gagagaaaag	tttaacccgcg	attcccatatg	125400
agggcccgagc	atccctctcc	cacctgacag	agaaaaccag	gaggacgcag	ctgaaaacact	125460
gtacaactat	aaatactcct	tcctagtgtg	gacagagttt	accaaagggt	atcgtaatct	125520
gaagcacaaa	cataacactg	taagttacca	gaagtttaga	cacgttgaga	atttaataga	125580
agttgagaca	tcaaacacac	gtttgatcct	aggggaatta	aattattggt	aacagaaaag	125640
ctctttacaa	agtctccaca	tttataaacc	aaatacactt	gtatcagaat	tgttgtctga	125700
gccagaggcg	gcggcacaa	ggtggctgag	gcagcaggat	caagtctgga	gaagcccctg	125760
ggtccaggga	gcactagtga	attcaagtaa	acattaaaca	ttaaataaga	aagacggatt	125820
ctacaggagc	ctagacaccc	tgtaaggcca	gtactactaa	gataccaaaa	ccaaactgac	125880
tcagggggct	gatgagatgg	ctcagtggtc	aagggcatgc	actgctcttg	cagaggacct	125940
gagttcagtt	cccagaaccc	gtatcaggca	gttcgcccac	ctgtaatgta	ccagagaccc	126000
taaacatttc	tatcctccaa	gcacagacac	acataaacat	aattaaaaat	aatccttaaa	126060
aaaaaatctc	tcgtggacag	acagcaaaaa	atactgaacg	aaatctaata	aggtagactc	126120
agcaaaagcaa	atccaaatat	gcgaaaagga	taatacaatg	caaataccaga	cttatcccag	126180
aaaccccgag	tcgcttttagc	atccaaaaaa	ttcaatcatc	ataattcacc	gtattagcta	126240
accgaaacag	aaaaggcatt	tgattattaa	taaatacagg	gaaagcattt	gacaacattg	126300
accattcact	cttaataaaaa	atgttaccaa	acaggagaaa	gaagtaaaat	ctcctcaacc	126360
cgatgaacag	gaacacaaa	tggctgggca	cagtggcaca	tcctgaaacc	ccagcaccca	126420
gatagctggc	agactgggct	ggcatgagaa	caaagggtaa	gaggcaccca	aacacttaat	126480
gatccaaggt	gctgcctccc	ccccccccc	gtgtcactta	aatccgtaat	taaatttctg	126540
ggaggagggg	tttgacacag	ggtctcactc	tgtggcacag	gctagcctga	atggagtcaa	126600
tgctggcctc	aaaactcttc	tgctcagtt	ccagagtgcg	gggaaacaat	catgagccac	126660
ctcaccaggt	tttaatgcca	tttaacagta	caaggttaaa	cgctgcctga	aaagaacaag	126720
acaaagaaag	atcacacaga	caaacatcac	acaggtgctg	tttgagacta	ggtctctagt	126780
gatgtaggct	ggtctcaaac	ttgcacaaga	aaagaatact	tttgccatca	agatcaaggt	126840
tgctatcacc	catggtggca	tttcggaagc	agaggcagga	agatcagtag	tacaagggtca	126900
tcctcagcta	ccatgggtgt	gaggccagcc	aaggcagcac	gtgagccagc	tgtggaagca	126960
cacatctgta	acccaccac	tcaggaggct	gaagcaggaa	gttcaagcca	cacaagagcc	127020
acttaaaaa	aaataaaaa	ataaataagg	ttggagagag	ggnnnnnnnn	nnnnnnnnnn	127080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127140
nnnnnnnnnn	nnnnnnnnnn	nnngaaaaca	cccattagac	tgctgaagag	actgagaaat	127200
tgctcatctg	aagggttgaa	atcactgagg	aatggctaata	gacaaaactg	cttaaggaga	127260
atttttattaa	aaaacaaaa	tccttctgga	ggagtactgt	gctcccttca	gcattaaacta	127320
ctgtgccagg	gaacccactg	ccagcagaca	cccaacttca	aggataatca	agcttgctac	127380
gtaacatggc	tgcttaacga	gacaccacac	acacttctctg	tatgctgggc	atcacctcta	127440
gcttacttat	aatatcaaac	ataagtaacca	tgtgagcgct	acagggcctg	gccaatgaac	127500
agcgatggga	aaaataacct	gcacatggtc	agtaatgggc	aatccactcc	ccagacactt	127560
ctaacctcag	ctgcagactt	gtgggtactg	agaacggagc	tgactaagtc	aattcacaga	127620
agcagagccc	tgtgcctgtg	ccagggaacg	gaggacgctg	ctcaatgggc	agagcttcag	127680
ggcggaaga	gatggaaagt	ggagctaggg	cccgatgacg	tgaatatggc	taatgccata	127740
ttgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	gtatgtgtat	gtgggtgtgt	tgtatgtgtg	127800
tgcatgtgtg	tgtgtgtgtg	tgtatgtatg	tgtgtgtatc	caaaatgggt	ataaaaaatcc	127860
catacaatgg	ctatgcttat	gtgtctttta	ccacaagtaa	aaaattttta	gtaatcctag	127920
gaacacattt	ctaaaatttg	gaatacttgt	tggggaagct	atctgagccc	cagcatgaat	127980
caggaaatgg	gtaggagagg	caggagagat	ggcttagtgg	ttaacacaca	catggtggcc	128040
ctttcctggc	accaactagg	tcactcacaa	ccaggatctg	acagcctctc	ctggcctcct	128100
caggcaccag	gcaggaaagc	ggccccacat	ccctgcaggt	aaaacatttg	tactgaaact	128160
aaataaaaa	ggagcagaag	ctgggccttg	aagtccttta	cctccagcac	cgggaagcag	128220
gtggatatatt	tgtgagttcc	gggtctacat	agtaaaaaact	tgctgcgaag	taaaacaaaa	128280
caaaaaactgg	gggctgggtg	gatggctcag	tgggtaagag	cacccgactg	ctcttccaaa	128340
ggtccaaagt	tcaaatccca	gcaaccacat	ggtggctcac	aacctccgc	aacaagatcc	128400
tcttctggag	tgtctgaaga	cagcaacagt	gtacttacat	atattaataa	ataaatcttt	128460
aaaaaaaaata	ccttttaaaaa	aacaaaacaa	aaactccac	taaaataatt	ggaaaagtcc	128520
agggaaagcc	acactcgatg	gcgcaagcgt	gtgatcacag	tattctgtag	gtaaggcaag	128580
aggatcacag	cgggacggag	gccagcctca	gctacacag	ccggctctggg	ctacagtgtg	128640
agaccccggt	ctcaaaaacaa	aacaaaacaa	aaaaagcggt	cacattatca	tatactcaag	128700
gccacaaaac	accttcttct	tcccaatcct	tgaatgttat	cgaatttggt	tacttttttt	128760
tggtaatattt	ttgtttgctt	ttcaaggcag	ggtttctctg	tgtagccctg	gctatcctgg	128820
aactcattct	gtagaccagg	ctggcctcta	actcacagag	atgtaattct	ttttcaagac	128880
tggatttttga	tttgggggtg	tgtgggtgat	accacaggga	cacttggggc	ccaaacccaa	128940
ccaaagaaaa	aacaccccc	ccccaaagta	tgaacatcaa	ctgtatataa	aactaacagt	129000

tcataataagc	taagtagcct	gcagggtacat	ggttatggaa	acagctttat	catacagact	129060
tctttcagta	gatgccattt	gagaaaaaaa	aaaaaacaaa	acattctttg	gattttcaac	129120
atgtaaacga	aaatatgtta	aatcttaaaa	aaccttaaa	cagacgccac	tgctctttgg	129180
ggtcaggga	gcacggctgc	aggcccccag	agcagcacat	tccctgaggg	aagccttggt	129240
cgtcctcacc	agcaccaaga	acagccaact	ataataagct	cataaaaaat	cctgaaaggc	129300
tgcccaagcc	tgaaaatctg	atatgaaaac	agagggcaag	acagataaaa	ggcaaacctat	129360
actttaaaag	tcaattccag	ctatctgtgt	ggaaggactt	ctcgggacgg	actatctgct	129420
aagaccttgt	gggtgattta	aacagcgtgg	agggcagaaa	aagaaaggaa	agtgcaggaa	129480
acaatggcaa	aagctcctcc	tccgtggcct	ctacacgcac	ccaagcgata	tgaggggagg	129540
aggggaaggac	actatacaca	ggtaccaatc	cccacgaact	agaaaacaca	gcttttacta	129600
actagtctat	tttttttttt	accttagcta	gtgtctttct	tatgtttggc	ataatttctg	129660
ctctgcattta	aataaaccta	gtgtataaga	aggcaaatga	gtaaaggtaa	agtcagttaa	129720
tgttcatttg	ttttgactgt	gtgggtgtgca	tgccaggcgc	attcgtggag	gtcagaggac	129780
aaccgcatgg	agcaggctct	ctccttccac	caccccagg	gtcgcagggg	tcccagggat	129840
ggaactcggg	tggtcagatt	tgcttgga	gggcttcaact	cactgaggcc	ccccagggt	129900
cccaattcac	ctgtttactc	taatgtatat	atttttgaga	cagggctctg	ctatgtagtc	129960
caggctagct	ttgaaactcac	tgttgacact	agctggcctg	gaggtctgag	tgaacctcct	130020
gtttccaccc	aagtgcgtgg	acaacagaca	gcaaccatca	aggcaaaatg	aagtcttccc	130080
ttcacaccaa	agtggcttcc	tatgaccttc	tcttgacaac	cccaaacagc	aagtgcctcg	130140
atatgcactt	gcgtttctct	ttttttactt	caagctttga	ctccactttc	ctaaaagggt	130200
tttactaggc	tgaagcacac	ttcaggagac	acctgtgagc	tctggagtga	ccggaaacac	130260
accagctctg	atgcaaaaac	aaaaaacagc	gatggggatg	gggcaagggt	gtgtgtgct	130320
tctgtgcaat	ggcattccct	gacgcattac	ctcacaccag	atacatacag	aaaagacaaa	130380
ggtaagtctt	cactcagtgc	gcgcattccac	cacacccaaa	gcaagcattc	aacagctcag	130440
tggtcataagg	ccagcattat	cacagcgcgg	acaaacacaa	agccaacctt	ctctttgagt	130500
tccccaaggt	acatagtgtt	ctgtccaagg	atggcctctg	cagactcctg	gaaacacgtc	130560
acccactgggt	tctcttgaaa	atctgcaata	tttgccataaa	aaacacaaact	ctattatttc	130620
agcgaataaaa	taacaaaagg	aagacacccc	aaggacgagg	gaagattatc	ttttcctact	130680
taaagtaact	tccaaggag	aaatgacatt	gccgccacac	acagcccccg	actctttgggt	130740
taacggttcc	cccttatgag	gaagtattca	cttctgggtca	ccccgtgttt	cacttttagt	130800
acagtatcta	atgaattaca	ggagccactc	aataccttta	tcataaaagt	ggctttgtgt	130860
taggtgatct	tgaccaagcg	taggctatgc	aaatgctctg	agcacactta	ggggaggctg	130920
ggcgtgcta	gccctgctac	atgcaggtaa	tgctacgcca	gtcttacgcc	acaggaaaga	130980
gcttaaaatg	tggtcgaaact	caatctgcta	gaggtgtgag	tttagacctc	gggagacggg	131040
ccaagtaaac	ttccgcagat	gtcggagcat	cagacagagc	tgtccctctc	tgatgcaaaa	131100
ggcttcgcac	ggcaagtttt	taatgagcct	ccatggtaat	agtgggttcc	tctctctcct	131160
cctcccttat	caatatgctt	gacatctttt	agttttttga	gacagcaaat	aacgtagccc	131220
aatctgccc	catacttact	ggcgagccac	agctggccct	gaactcctcc	tgctctcct	131280
ttaccccgta	acaagtgtctg	gggttacaga	catatgccgc	catgttcagt	gtaagtgacc	131340
gctgcccagc	agggcaaggt	cttccttatt	tacaaagcag	cagccaagcc	ctgcagccca	131400
ggcctatctg	atttcctcag	cacaccccca	agggtctcac	cgatggcagt	cagtccatga	131460
acaccgtagg	cttctcacaa	tccacactac	tcaccgataa	gatcatgcgg	tatttgaaat	131520
tgggaaattc	ccggtcacac	ttctcacagc	ggtacaaacc	attctgctgg	tcaatcactt	131580
tcttattgca	gtcctgggtt	gggcaggcct	ggtacataca	gttctctttg	cggagaaaca	131640
ccacogctgc	cacagtgtctg	aaatagtcgg	cctgcaggcg	aaaggaagac	cgccatcagc	131700
aagcaacaca	ggtctggaac	aggcaactca	aagctgctct	tccttgcaag	tggtgagcgc	131760
gtgtacatcc	tagctcccac	gctcacgagc	gatatgcaga	atcactaact	ctgggttccag	131820
aaatcacacg	tgctacacgc	agaacccaaa	gaagtaacaa	accggcactc	acgcacctca	131880
ctttttccat	tttgagagcg	gcggctcact	agctagcctt	gaactcagag	tttggctctgc	131940
ctctgtctct	atagagtgcc	tggtctaccac	acctgggcac	tttgttttcg	agacagggtt	132000
tctccggaac	tcactctgta	gaccaggctg	gcctcaaaact	cagaaatcca	cctgcctctg	132060
cctcccaagc	gctaggatta	aaggcatgtg	ccatcagcgc	ctgaccaat	tctttttatt	132120
tatagtattt	atttggtaca	tatgtgtatg	agtgtgtatg	acgtccatat	gtatatgact	132180
gtgtgcagtg	tgccatgtg	tgccaggtcag	aggacaactc	tcaggagtca	gttctctcct	132240
cctactgtgg	cgtttgggga	actcagggttc	caagatagca	ggaaaagtgc	ctttaacagc	132300
tgagttatct	cgacactgac	atgtaatgat	tcagtgtgca	cacagtgtt	ttgcatgtag	132360
tcataataac	cacagcgtgc	atgtgcaggc	taggaaacaa	cctgtgggag	ttggtgatct	132420
ccctcaccac	tgtgagtgg	ggagaggaa	tcaggctgtc	aggctcgtg	tggtgcctt	132480
tactcactga	gttcccttgc	tgcccaagca	ttatttataa	gatggtgatg	tctaccctta	132540
tctttaggga	tcagtaagtt	tttccccaag	acaagccaga	aaaacctttt	aggcccaatg	132600
agccattcag	tcagtctcta	ctgccactcc	tcaacctgtc	tgtggggcag	gacaagccac	132660
agacaacctg	aagacggaag	gtgagccaat	aaaatcttac	tgacagtaac	agccagccag	132720
ctcacaggct	tgacaggcaa	ttcttgact	aagatgcgtt	aagagaatgc	agaattaccc	132780
atgactaaga	tcttctaata	ggaaaaatgt	ctggttaagt	aaccagcaa	ggagctaagt	132840

cacgcaagcg	gtggatacct	gctttgcctc	tgaaccctgc	acagggttttg	gtttatgttc	132900
aatcatgtca	agtacctaca	aaatccagat	ctagcctgaa	ttcaaagata	ggctacctac	132960
cctgcecccg	gacccccacc	cggggtctca	ctgtgtagtc	ctggctgtcg	tagcgctctc	133020
tatacagacc	agctgttatc	aaattcagag	accacttgct	ctcccaagtg	ttgggattaa	133080
aggcatttgc	cactatgcct	ggctctcact	ggttgtagca	gaggagcaaa	acaatgtggc	133140
catttaaaga	gacgaagcta	gaaaccagtt	tagacagctt	tggggctgtg	ggtgtagctc	133200
agtgaagag	cttgcttagc	atgcacaagc	tgttggtttt	aaccctcagc	gtgacagaac	133260
cgaatacata	agaaggctta	gaggaggggg	tgggtgtgga	gagatgggtt	agaggttaag	133320
agctgggtgc	ttaatctccc	agtccccaca	tgggtggctc	caacatccat	aactgcagtt	133380
ccaggggatc	tgatgtcctc	ttctgacctc	cttggggaca	tgcgactcat	ttggctcaca	133440
tgcaggatgc	ctcggggcac	tatgtggctc	caggggtgaa	ggtgcttgct	gccaagctgg	133500
gtggatttga	gtttggtccc	tgggacctac	aaggaagagt	ttgacttcta	tacactgagg	133560
tggaaacatg	atgtctctac	cgcaaattaa	aaacttaaaa	tttaaagagg	aagctgtaga	133620
gaaatagctt	ttaggaggat	gcctaaggaa	cttctctgct	ttttcagggt	agattcagac	133680
tcaaagccca	atttaaaagt	ttgagtgtct	tcacgtgttc	tgtatgcccc	gttctggctg	133740
tccgttgtct	gtcttttagtt	taagagagca	actgggtgag	aagtaactga	gagtctagcc	133800
gatgtttaat	tctcaagatg	tcctgtgata	gcattataag	ttgctgtgga	tgacagtgat	133860
ggatcagaca	cacagtgaac	tgagacagt	aggaagaaaa	cctatatatt	atacaggagg	133920
aataatggag	cagtggcgag	cgatatattt	ggaatacaat	agaaataaat	gattcaataa	133980
tccaagagaa	cgcagggttag	accaaggtag	gaagaggatc	actgggctgg	agagatggct	134040
cagtgtgtta	gagcactgac	tgtctctctg	aaggtcatga	gttcaaatcc	cagcaaccac	134100
gtggtggctc	acaaccatcc	gtaataatat	ctgatgccct	cttctggagt	gtctgaagac	134160
ggctacagtg	tacttacata	taataaataa	ataaatcttt	aattaaaaaa	aaaaaaaaaa	134220
aggaagagga	ccactgagca	cacattgaaa	tggaagatgc	aactgaaatg	caataccaga	134280
gccagtctgt	ggcatggcag	caggaattac	agtctgttca	aaaccagca	cggacctgag	134340
gctacattat	gtccctctac	actggccgtg	actgaaaagc	agcaacgtgg	taccctggag	134400
caggcctgag	gtcccaggta	aggacacagc	cctgacctgg	cactaggaaa	caggtgtaaa	134460
aactcaagcc	cctgccgttc	tctgggtgtc	tggagcaggg	ggtgcgaggt	acctgtctc	134520
cctggcccg	gttctcagat	ttagcctcat	gcaaagtttt	ccagttgggt	ttgccccctc	134580
cgccccctcc	actcctgtgg	tcagagatgg	aaacaccatc	taaggcttgt	ccttctgagt	134640
caaacctagg	ggagaaagaa	acaaacgtac	tgctgacagg	aacttcacat	ccttctcaaa	134700
tgtgtgtgta	atgcaccagc	cccagctgct	gccccctcgg	ctctcaccta	ggtctcagca	134760
cacctagtcc	actcaagaaa	tgcaatgccg	tgtctccttc	catctcgtgt	ccttcaatgc	134820
ccctcctcct	ccaccgcgaa	ggtcagacaa	actgccaaag	taccacagac	cctttcccta	134880
aactgggctc	acattgacag	cagccatgaa	caaaggcagc	aacagcaagc	cacactggca	134940
cccctcccc	caccctcggc	ctcaagcctt	caccccaatc	acaaaaatga	accagatgaa	135000
aagggaacat	tttgtttcat	gagtctgcca	aaaaatgttt	catgggagaa	aatcttttaa	135060
gcccaatcag	tatatagaat	caatcagaag	gcattatctg	ctgtgttaga	gatttgcatt	135120
ctcagtgaga	atttgtttat	ggaaatatca	aggttaaatt	cattttatat	aaattattat	135180
taaaattgac	tatacttctt	gacctcttat	tggctaattc	gaaggaaaag	aggccaggga	135240
gaaggtagca	aggacaggct	agtggcagag	gacagtgagc	ctgagatgaa	aacacctcaa	135300
ctacacgatt	ggtggttaact	aaagctgccc	ccacacagga	ctgactccct	aaggagactc	135360
ccaacagagt	agcgggtggc	caaggcaagc	agtcacatcc	gtcagaggac	aagggtccata	135420
aatccgtctc	accactacta	gccgtcgaag	tttaacttca	cacaggcatc	ggcaacccaa	135480
gacggctcat	caggagtga	cagagccagg	agacaggcag	cctctcttat	agtaaaaggt	135540
ccactcttag	agccaagatc	taaagatcgc	taatccttgc	ctgggtaggg	gggggtatact	135600
gaacttgtag	atcaatgaca	gtattttgtt	cgggggacag	acatgcctgt	ttgtgatttc	135660
tgtgatcaca	gttactgaca	gtgaaatatg	aacaccttaa	tgtagacaca	caattaacct	135720
ggggttgtgc	tcagagtctg	ggagaaaaac	caaccaacaa	accagaccaa	tgccatggct	135780
gtctctggtc	agacttctgt	cctacaatgc	aacgctagct	atcagcacag	ttacacagcc	135840
agaccacac	cactgaggat	gccttaagt	acagacaccc	cagatggatg	ctctaagtag	135900
taaatattgt	tattggttta	catcagactc	aaaacaaaaa	tgaggagcgt	gaaatatgag	135960
ctgtttttct	tgagtctctt	aaattctacg	acacagctgg	aaaccacaca	tgccccacct	136020
ggtactacag	ctattcaatt	tccttaagct	tgggctaggt	aaagtatttt	ctttgaccac	136080
ctcaatcctt	ctacacaaca	ccctcaggat	gaagtgtctc	tccaaggcag	agtttaaatg	136140
cttaaaaagc	tagatatggt	gtggaggttc	taacaagctc	tcctcatctc	ctgagtgttc	136200
tgggaatgac	aatgtacagg	caacaaagag	tccattgttc	tcactgtact	ttaccgagtc	136260
cagtgaacac	acaacagaag	atgtgaccag	ctccagagac	tcagcagagt	aagaaagtac	136320
attctacccc	cacatcagct	gcccacagt	caaactgaag	atgtctctcag	cactgttctt	136380
cagcgcgcag	ctgttaaagc	cttgccgtac	ctggcggtat	caaggtcagc	acttagattc	136440
taattgcttt	ctctggctct	ctatgttcat	caggggatcg	tgggtttgtt	ggtttaagag	136500
ggctttcagg	ccaccatctc	agaagacata	atgcctcgaa	gaagtatatg	aaccactcca	136560
ctcaatagca	aaggcttgcc	aaaacagctt	ggcagcagtc	cacatgctga	cagtgtccag	136620
ctctggttta	ccagagcccc	agcacagata	acccctgagc	tgagttgcac	aaaacctacc	136680

agccacgaag	cttataggcc	tctgggatgt	caggattcac	aatgacagt	ctggatgaga	136740
ggaccgagag	gctccgtcca	ccgaagtcag	agactcgggc	tcctttgatg	gccatcacgg	136800
gctgccgaga	gccgtcaaac	ttgtcagcct	gcaaatcaaa	cgcacgcaca	catcactgac	136860
aagaagaatg	tacaggatgg	ttttatggag	aagagtcagt	caatgtctgt	ggactctaag	136920
acagacctcc	cactcgggaa	ggagcattgc	actacaagaa	gctgcaataa	ccgatcatct	136980
cacacagcga	aggtcttcaa	atacttactg	gatagcacia	ctccctgagc	taaagcccca	137040
ccctcaggac	tcggctcagt	gcaaggactc	acatcttctc	cccacagagt	tgtggtcacc	137100
accttccctg	acatgtccat	caaatagata	tttctcttag	caacttctct	gttggtcgac	137160
ttcacctgtga	ttttaatcga	atcttcatag	ctcttgacga	ttccaatgat	gtctgaaaca	137220
aagaacactt	tgtaagagct	cccatcaagg	ctcctcttta	aaacacggca	gggacgaaag	137280
gcaagcacgc	tagtcacgca	tcaccccaaa	cactggagac	aaaaatcacc	actctttgcc	137340
ctcaaccttg	gacgcaccta	ctagtgcgtc	tttagccttg	ctctctaggt	caccgatccc	137400
tgtgaaatca	aactgaactg	tgggtaagtg	atggccatct	tcacagggaa	ggacagaagt	137460
ctcattattg	aaggtcatct	catagtcatt	tttaacagcg	gagaactggt	tgttagcgat	137520
cttcagggcg	ccctttgaga	agtaatacac	ctgcaaaaaga	ggccaagtca	gggcaagcta	137580
ctcagtcctt	caaagcccg	cccacaatgc	agccttccaa	actgtgtctc	tggctctcac	137640
agttcagcct	gtcagctctg	actcaacacc	ctcttcttac	cttgttcaact	tcaataaggg	137700
gaaagaactt	gtccacttgc	tcattgaaag	cagtagctct	gatttcaccc	tgcagaagca	137760
agggagagca	cattagaact	gcgctgctag	gccctcgctc	tcaacagcga	gagcaggcca	137820
ttttgacagt	tagacaccat	cctcggagca	acagtcctaa	tgtaaagaga	ccctcacagg	137880
gctgcagaca	ttgcagaatc	ccacatatat	agttacattt	tcaagctatc	agctctggta	137940
aaacaaaagc	agctcagcca	cccgagccta	cagttgttaag	ttaacacata	aaacgaggga	138000
gcctcgagat	gatctggaag	acaaaggtgt	ttcctgtgcc	aacctggcca	ccctagttta	138060
atcccggaac	cctcgggatg	gagaagacaa	tcgccccacc	aagttattct	ccagcctttg	138120
cacacgtgtg	catgcataca	cacgaaaata	attctttacat	cttatcaaaa	cattttttaga	138180
aggaatagct	taatatcccg	ataatgaaac	ctaattattct	gggtcccag	tgtgggctg	138240
gaagcccaca	ggaaggcggt	ctgaactcca	atgattcagc	actcttcctc	acgcccgcga	138300
cctaagtgtac	ggcttacttc	acagtcagga	gctgggtgcca	aaaatgttcc	aacaagtcac	138360
gctttcaagc	tgactgcact	gttacttttc	tgtcaggcaa	tatgttacag	atggataaga	138420
aacttaacta	aatggctaaa	accttaccct	gacaagctgc	aggaaattag	tttctgatac	138480
tgaatctgcc	atcacagtta	agatatctgc	tgtgtctacac	cgtgacaaga	ggaagaagag	138540
ggggagagag	ccccaaagct	ttgccccttc	cggtatcatg	gcttattttta	cgtgtgtggg	138600
tgtttttacca	gcatgtatgt	gtgagcatca	ctgagtgctc	gatgcctgag	gtgaccagga	138660
gaggtactgg	atcactgaca	cgggagtcac	agatggttat	gtgctactct	gttggtgctg	138720
agaattgagc	ccaagtccct	tcaaagaaca	gcgagtgctc	ttaaatgctg	agcagctctc	138780
ctggcccctg	catttttgctt	taagtaaagc	ctagttagta	taaatagatca	gaagccctcc	138840
ccgcagacta	gttaaagctg	agtagctgct	gctcctctct	ctggaggcaa	acccgcccctg	138900
ctcggcagggt	atcaccccag	ggcttcaact	gtgcccctag	caatgtaatt	agagcgctgc	138960
agtctctgca	gacggagact	atttacagtc	caccaatctg	tattgctaca	gacgcgtccc	139020
tttaggagca	ctttatctca	tgtctgaccc	tgtgccccag	atagcatcaa	aggtctccaa	139080
cagaaggaaa	gccacaatga	ccgccatgtt	cctcggagct	gagccccacc	cgnnnnnnnnn	139140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139260
atctctatat	atgaatgagt	atcctgtagt	tatcttcaga	cacacagaa	gagggcatca	139320
aatcccat	acaggtggtg	gtgagccatc	atgtaattgc	tggaatttaa	actcaggacc	139380
tctagaagag	caatcagctg	tcctaaacccc	tgagccccct	cccttttttt	tttttttttt	139440
ggcagcttat	ttttaattgt	tttaaatcac	gtatattgtt	ctgtatgtag	ctatgttcac	139500
ataagcgcag	gcactcatag	aggtcagaga	tgtaagatgc	ctttggagct	ggacttaaac	139560
atggttgtga	gccatctgat	atgagcacca	agtggacttg	ggctcctctg	accgccttgc	139620
tgccctccac	ctagaattac	agtatatgca	attagctgct	cagccacccc	ctcagtcctc	139680
cttagtactt	tcaacatagt	gtaaatcgga	gagcactcct	tggtgcaaag	attctgtgtg	139740
ttctatcatt	ctctacaccc	aggcttcatt	gggacgtgtg	atgccagcga	tgattcattg	139800
gcgtgaagta	tgagtttagag	tcagacaatc	agtgctctct	gctctgagtg	ggctctccctg	139860
cccctgtgtc	tctgagacag	gggaggatcc	tcctgttgga	tcacatagtg	catagagccc	139920
tgtgtgcaca	gcgctctcaa	ccatttgttg	ccatttcaaa	tacagtgact	tcagagtctt	139980
ccttgaaata	acagatttct	ccatttgttg	tgttcccttc	tttctgcaca	aagacctgag	140040
gctgagatgt	ttttgaagga	actttctagt	agttactcgg	tggaagagga	caatgatgct	140100
cccctcttct	acagagaaga	aggaacagg	aaattccaag	aaaaggagta	cacagatgct	140160
gcagtgtgtg	actctaaggt	aacgtgcgtc	caaccagcag	ttgaacaagg	cggaagatag	140220
aggcaaagtg	cagactcata	accagtccat	tgttttgctt	tggttagttg	ggttgggttg	140280
tttgtttgtt	tggttgtttc	ttcttcttta	ttagtccgat	tggtattggg	aaaaatgtag	140340
tagaattcta	ttgttgatta	tattcacaaa	caaaagactg	tttatatgtc	ctgaattttg	140400
gcaagattct	ttccatattc	cttcagacc	atagcagcag	caagcttagt	ggccctttgc	140460
caggatttct	cactgagctg	tacagcattt	gtctgaagaa	ctgatgaaca	tttttagctta	140520

tgtcctctatg	ttagtaaata	gctagggtttt	tagctagcca	ccttgctagc	ttacaactag	140580
aatgcctctg	gctgagtttg	gtggcacttg	ccttttagttc	cagcactcag	gaggcagaag	140640
caggtgaatc	tctgagttca	aggcagcctg	gtctacataa	caagttccag	cctggccagg	140700
gatatgtagt	aagactctgt	ctcaataaag	taaacagcca	gaatgactta	aatattgcta	140760
aaaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	agaaagaaag	aaagaaagaa	140820
agaaagaaga	aaatgagtgg	atlttgccata	gcctagtcaa	actagatttt	cctgggtata	140880
tgtaacata	ctactattaa	caacaacaaa	tactttacac	taacatgac	aagctatatt	140940
tttaaatat	tattttatat	gtatggtgtg	ggaagctgtt	gcagagtggc	agttggctac	141000
tgctggccac	cacacataca	taggcagtga	aggttctttt	gccaagaca	gttaaccaat	141060
cagatgtgag	acacgcctct	cctaggcccta	tgtaagcagc	accagttctg	ggctcagggt	141120
ctcttcgcct	ctacaatcaa	gctctcccaa	taaacgtgtg	cagaaggatc	ctgttgacgc	141180
gtcgttcttc	ctggccaggt	gagcgcgcac	aagagtattg	gttttcacct	acatttatgt	141240
gaactgcacg	tgcaacttgg	acccaggagg	cccagaagag	ggcatcatat	cccctggaac	141300
tagagtccaca	ggttatgtat	gggttctagg	aatcaaacc	atgtcctctg	gaagaacagc	141360
cagtgttctg	gatttaactg	ctgagtcac	tctccagccc	cacatgatg	gatttaactg	141420
ctgagtcac	tctccagccc	caccaagctg	gatttaactg	ctgagtcac	tctccagccc	141480
caccacgctg	gatttaactg	ctgagtcac	tctccagccc	caccacgctg	gatttaactg	141540
ctgagtcac	tctccagccc	caccacgctg	gatttggaag	cagaactgag	gctttgaaca	141600
ctgtgttatt	ctaactcttg	cttctcttga	accctgagaa	aattccttct	attggccttt	141660
cagagatctg	tatgggctta	aaacaagaat	atgtccaact	cctgatctct	gattttatat	141720
attaaaagaa	tagagtgcag	ctaggagttt	ggcacactcc	tatgttccaa	gcttcaggaa	141780
gggaagaagg	tcaggagttg	aaggctagtc	agtcttagct	ttgtgacaag	tttgaggcta	141840
acgtgagctc	tatgagaccc	tgcttaggag	aggagaaaga	aaggggtaaa	gtgaggtgtg	141900
gtccacgct	accacttgcc	taacacgcac	gacgcccag	gtttccgttc	ccagccctgg	141960
ggttcagagt	attgtgcaca	ccctgtatct	ttaaaagtca	atagtcaacc	ccttggaatta	142020
actttccgaa	aaatattaaa	gctacatcat	ccactctaca	aacttgtaaa	agcccttctt	142080
tgtaatgggt	taagaaagta	tttggtctca	gttttatggc	ctttgcatcc	taaatgttac	142140
cccattggaaa	gttgcttaat	gaaaacaggt	aaaaataaga	caggaggagc	tgcaaagatg	142200
ctcaggctca	cgagtgccta	ctgcccctgc	agagaccatg	gcaggcagtg	ggctcccta	142260
ctgcagggaa	tccaaaaccc	tcttttgccc	tctgcaggca	accacattaa	cacatgcaca	142320
catacacata	atltttaaaaa	taaaaataaat	cttttaaaat	gagctctaga	acgagtttga	142380
catcagtcta	ggctacacaa	gaccttgtct	caagaagaaa	gaaatgaagg	tttgctgtgg	142440
atggagaggg	agatgcactt	cctattccat	gaagctgcta	ttttgggtat	tatggtactt	142500
tgcaatttta	tagagagctg	ctattttctt	tcttttaaa	taatgcttgt	tgttttgact	142560
gtaaaagtaa	taaatgttac	tttggaataat	atagagaagt	ataaagagta	aaaaaaaaag	142620
tcataaccag	tgaagtaccg	ttaacatttc	tgcttctatc	cggccagcca	gagtttttcc	142680
tgtagagaatg	tgtttttctt	ttacaaaatt	gggataatgc	tgcaacttact	gtttgttagc	142740
ccactctttc	ccttcacgat	ttattgtacc	cttttctca	cagtattaaa	ttttcagctc	142800
caagtaattt	tcocatagcta	actctgtgtt	ctgttacaga	gaagaaatgt	acttaattta	142860
agatctaata	ttggcataat	atlttggtcatt	atgatgctat	aataagcatc	tttctgtata	142920
aatatttttta	tgtacagcca	gtgttttgtt	tggtgatgaat	tagcacaaga	gtaaagtgtg	142980
gggtcaagcc	tctgagtgc	gctgaattgt	tctcaggaaa	ggactagtta	acatccattc	143040
tgaagaatg	tggtgaatgt	catttcttaa	gcaacactgg	ttattattac	atactattat	143100
tttggttagta	ttatacatctt	atlttaggggg	ataggaaata	tgctctatata	gttcaatttc	143160
aagagttgca	aaaatatatg	gtgtgggtttc	ttttcctgtc	ccctagttta	gtctcacttc	143220
cagccccata	atctaccttg	ttaaatatatc	tgtatataag	gcataatgtag	gtgaataaaa	143280
aacagcctag	tatggtggcg	ttaacctcgg	gttacttgag	aactgctctt	gcagagaaca	143340
tgactttggc	tccgaaagcc	ttcctggaat	tccagctcca	agggatgcag	tgccctcggc	143400
atccttgggt	acgtcactca	tgtagcacaca	tacacatttg	gttttttaac	ttaggaactc	143460
caagtggggc	aatgagatgg	ctccatgtat	aaaggcagtt	atgcaaagg	ctggatgaca	143520
tgagttcagt	cttcagattc	tgcaagataa	caggagagga	ccaacccctg	cgagttgtcc	143580
tctgacctca	gtacacatgt	catggtacgt	gtgtagtatg	cacatgcaca	gaagtcccag	143640
cactcgggag	gcagaggcag	gaggatctct	gagtttttagg	ccagcctggg	ctacaaaacg	143700
atlttagctt	atataaagaa	actctgtttt	gaaaaacaaa	acagggggtg	gggatttagc	143760
tcagtggtag	agcgtctgccc	tagcaagcgc	aaggccctgg	gttcagtcct	caactctgga	143820
agagagagag	agagagggag	agggaaagg	agagggagag	ggagagagag	gaaaggaagg	143880
aaggaaggaa	ggaagggaag	aaggaaggaa	gaaagaaaga	aagaaagaaa	gggaaagaaa	143940
gaaagaaaga	aagagagaaa	gaaagaaagg	gaaagaaaga	aaggggaagg	aagaaagaaa	144000
gaaagacaaa	gcaaagcaaa	actaaataaa	atacatacaa	tatatataat	tttaagactt	144060
gagggcatag	atcagtgtta	taatgcttac	ctagcatgcg	taaaactctt	ggcttctaaa	144120
cctagcacc	taccgcaaaa	tatttgcctc	gtcttgctaa	attatattgc	tagttgtcag	144180
actactgtgt	actttcacta	gcaacataat	gagaatgttc	actatcccac	tcctctgtca	144240
agtaatctgt	tcttggtttt	atltttcttt	gccaaattga	tggttgaaaca	agttatttcag	144300
ctagcctaga	acacacagag	aactctcttt	atgaactcaa	gtttcttatc	cttttatgat	144360

ctccagaggt	ttgtttttgt	gggtttatta	gtgttttttt	tttggttggt	tggttggttg	144420
tttggttggt	tggttggttg	tgctttactt	tttcttattc	atttttttat	ttctttattt	144480
ttttgttttt	aatttaattg	actggttcca	tgtggccaag	gatagcctca	actttgtagc	144540
agaaactggc	tttgaacttc	tgggtcttcct	tcatctacct	cccaagtgat	gggattaagg	144600
cacgtgccac	cacatctaac	aatatctggg	tttctttatt	ggagtttgaa	agggattcct	144660
ccagcattac	tttgactcct	catagtttct	tccagagtta	ttacactttc	atttgttaca	144720
ttaagagttt	gatccagggc	tggagagatg	gctcagtggc	taagagcacc	aactgctcct	144780
ccagaggttc	tgagttcaat	tcccagcaac	cacatggtgg	ctcacaatcg	tctgtaatgg	144840
gatctgatgc	cctcttctgg	tgtgtctgaa	gacagctaca	gtgtaatcat	ataaataaaa	144900
taaataattc	tttaaaaaaa	aagagtttga	tccattttaca	ctggacttct	tgaggcagca	144960
ggatcataaa	ttcaagggtga	gcctgggtga	actggcagaa	gtggcagaag	gtgtgtctca	145020
cactgctgta	ttcatttccct	cattgatcct	cagaggtttg	ctaacgggaa	gtaagtggaa	145080
cagaaggttc	agtattcctt	ttttcccaat	tctattcagt	ctttagtagt	agatccctca	145140
ttatctgaga	tgcagagtcc	cctttattcc	tgtgaccatc	tcgttggttt	tcagggagtg	145200
tctcattcaa	ggcctaacac	tgaggacatt	tcaactgtgt	atgccaatcg	ctctgcagcg	145260
ctcttccatc	tgggtcagta	tgaagtgagt	attgaagaac	ctggtgtcct	gcctgtggct	145320
gcagtggaag	atgagctcct	ctctgttctt	ctgcacacat	tgaaatcaac	tagcttgcaa	145380
acactgacat	ccaccagac	ccattctctc	ttctgactca	tgtcacctct	cataggtgac	145440
cacaaacaat	atgtagttag	caagaagtag	ctatgtcatt	gtccacagtg	catggatttg	145500
ttccaatagg	ccagcacttc	tgtgtccata	tcagctagat	gtgctgctga	tagtatttta	145560
gattccaaaa	tgtgtccaga	tattacctcc	ttcgttttgt	tcttctaaat	aagccaggca	145620
caaagacttg	aaagatggct	tgatggctct	tccagaggtc	gggtttgatt	cccagcccca	145680
acatagcagc	tcacaatagg	ctataacgtc	atttccaagg	ggtctgactt	cctgttctgg	145740
cctctacagg	cagaaagcac	agacatacat	gcaggcaaaa	cacataaaca	taattgaaag	145800
aagatattaa	ataatagccc	acgcttgagc	ttattcctct	gatgatacag	ctctcctgaa	145860
gtcatcatgg	gcagtgtaaa	agtaaagggt	ccccgccctg	cccaggggca	tcagtggagt	145920
agctgactgt	gagtggcctt	cttctcatcc	cctaactgct	gcaacatcat	caactgtgga	145980
gcattatata	cgtggatttt	aatgttaggaa	atgacaaaga	ttagatctat	ggccaggcac	146040
tagtgcacgc	ctttaatccc	agcactcagg	aagcagaagt	agggtggatct	ctgtcagttt	146100
gagtttacag	agagtgtcta	ggcagccagg	gctatataga	gaaattctat	ctggaaagaa	146160
taaacaaatc	agatctgtat	ttcaggaaga	tgcctatgag	ctgtttgaca	tgtgtgatag	146220
aggtccttaa	ggacaggaaa	gtattccaca	tgtgtgcatt	ttacagaaat	tggttatata	146280
ctggagttaa	ggactgttgg	aatagttgaa	tgttcacact	cagtgttctt	caaatcaaat	146340
agaagaacaa	gaatctatct	gggcatggtg	atgcacaact	gtattcctaa	catgtagaag	146400
actgaggcat	gctatgtgtt	tgtggctaac	ctgggctaca	tagtcaatat	tggaacagtc	146460
gagctgctac	taaaacctaa	aaaacaaaaa	tctaataatc	tgggatttta	tattttcctt	146520
tttttaaaaa	aagagtgggc	agaatgtctc	tgatttttgt	cagatggcca	cagaacctag	146580
aaaaactgct	gctgtcgtcg	ctgctgcata	gcacacagct	aatattttgac	tatacgtata	146640
taaattttgt	tgtataactt	agctgtgctg	tcaactttgg	aaaaaaagta	tcccagttta	146700
tcattttaaa	ttggcactgt	acagaaatta	acagccatat	tagtctagac	acattaaact	146760
tcatttttoc	atttatacag	aagcaatgta	ctgtattaaa	tattcagctt	tatctacagg	146820
ggtttgatta	cagaaactat	caaagtattc	tctaattgat	gaaaaaagat	tcaagaatct	146880
gactgtagat	ccaaaggaca	agtggagaaa	aactttaggaa	gaattttccc	tttatccctt	146940
ccctatattg	atcatctctt	ttacttctaa	taatagtggc	cattttattga	acatacccag	147000
gagttccttt	catcacttta	gatataataa	ttatctcatc	ctaaaatgac	ctgttgatga	147060
gtcatctctc	tttcagatga	gaaacaaaga	cattgaaaaa	tctaacttgc	ctgcataaga	147120
tcacacctag	cctcttactc	actccatgaa	tattcctttt	ttttttttcc	tttgagacag	147180
aatctcacta	tgtagtctctg	gttgtcctag	aactcaatat	atagaccagg	ctagcctcaa	147240
actcacagag	atctgatagc	ctctgcctgc	cgagtgtctg	ggttaaatgg	atgtgtcacc	147300
aagcccagca	aaattttacct	tcttaatttt	ctaagacgtt	tctcctctta	aaaaatggaa	147360
ctattgagcc	agtcagtgtg	agacaggctt	aatctttaat	ctcagcactt	aggaggcaaa	147420
gacaggccta	tgggttcggg	gcagcctgat	ctatagagag	agttctatgg	gttaaagttt	147480
agggttaaa	ttttgagaca	aaactttgtg	tcaaaaacaa	acaaacaaag	ccagactgct	147540
taataagaca	aatcagacat	aatattataa	acaagtatta	gtgtcactca	attaaaaagt	147600
cactcaggag	gtgagatcaa	tccccagcat	aatgagggga	ggagggggaga	gaaggaaatg	147660
aatgggaggg	gaggaggaag	gaagagagaa	aaaggaaaga	tctcaaagca	gaacacagga	147720
tgtaatttaa	ggcctaagct	ctcgactgaa	gttgtccact	tttaatgacc	cttttcatgc	147780
tcactggttt	ctgtcttcgg	tacatagtga	agagtggaaa	ccaaggggtc	tcctgtgaat	147840
tecttttgtt	ttcaggcatg	tcttaaaagc	atagtggaa	caggtatgca	tgggtatcct	147900
gaaagactgc	agcccaagat	gatgggtgct	aagacagaat	gcctgggtgaa	cctggggaga	147960
ctccaggagg	caagacagac	catcagtgat	ctcgaagca	gcctcactgc	caagccaacc	148020
ctggtgcttt	cctcttacca	gattctgcaa	aggaatgtcc	agcatctgaa	aataaagatc	148080
caagaaaagg	agactctccc	agaacccatc	cctgcagctc	tcaccaatgc	cttcagggat	148140
atagccctgg	gggaagagaa	cacacagatt	tctggggcct	ccctctctgt	cagcttatgc	148200

acacaccctt	tgaaggccg	ccatctagtt	gccacaaaag	acattctccc	aggagaactg	148260
ctggtgaagg	aagatgcttt	tgtaatgttc	cttatcccag	gagaaatgcc	acgacctcat	148320
cattgccttg	agaacaagt	ggataccaga	gttaccagt	gagacctcta	ctgtcaccga	148380
tgtctgaagc	acactttggc	cacagtacct	tgtggcagct	gcagctatgc	caagtattgc	148440
agccaggaat	gtatgcagca	ggcatgggac	ctctaccata	gcacagagt	ttctcttggg	148500
gggctgctcc	tcacactcgg	ggtcttctgc	catgttgccc	tgagaatgac	tcttttagcc	148560
agatttgaag	atggtgatag	agttgtaagg	atgctttgtg	acgaggttgg	tagcacagac	148620
acctgtttac	ctgaaagcaa	gaatctggtc	aaggcatttg	attacacaag	tcagggagag	148680
agtgaagaga	agagcaagat	aggtgaaccc	ccaattcctg	gatgcaatgt	caatggaaag	148740
tatggaagta	attataatgc	tatcttcagc	cttttgcccc	atactgaaaa	gcatagccca	148800
gaacacagat	tcactctgtc	catcagtgtc	tccgcactgt	gcagacaact	caaagctgac	148860
agcgtgcagg	cccaaacctt	aaagtccctt	aagctgaaag	cagtgacccc	agggctgtgt	148920
gcagatttga	ctgtttgggg	agcagccatg	ctgcgacaca	tgctacagct	gcagtgtaat	148980
gcccaggcaa	taacatccat	atgtcacaca	ggtaagtcag	aaatggtttt	tacttacatt	149040
attggatatt	caagagctaa	tgtttaagga	gaaaaacact	ataaaggaag	cctggcatca	149100
aataaatcag	tgacctaaaa	ggaaaaacaca	gccgtcttat	atatcattat	gctattgaga	149160
agctttgagc	acatttctgt	gaaccagag	cttgggaggt	ggagatagga	tgattaggag	149220
tctaagacag	cttttagctat	acagcacgtt	tgaggtcagc	ctgaactaca	tgagaacttg	149280
tctcttaaaa	acttgagcca	gagccaggtg	gtgggtggc	atgcatttaa	ttctagtact	149340
caagaggcaa	aggcagggcag	atccctgaat	ccagcctcgt	ctatatagtg	agatccccac	149400
caggctacat	agtaagatcc	tgtttcaaat	aaataaata	aacaaaaaca	gcaataataa	149460
caatagcaac	aaattaattt	tttagatgta	ttattttatt	ttatgtatga	gtacaccatt	149520
gctttttcca	gacacaccag	aagagggcat	tggatcccat	tacagatggg	tgtgagccac	149580
catgtatgtg	gttgctggga	attgaactca	acacctctgg	aagagcagtc	ggtgctctta	149640
gccactgagc	catctctcca	gtccattaat	taaaaattta	aaactagagt	atttttaaac	149700
atatttcat	tttgtgtgtg	gtatacatat	tataatacag	gttcataagt	caattctctt	149760
ctaccatgtg	tgtcttgag	atcaaaacta	ggttcttagg	catgggagca	agtatttact	149820
tcctgaacca	tctccctagc	catttctagt	attcttttct	tttgtcttga	aagatttatt	149880
tattatatgt	aagtacactg	tagctgcctt	caaataccgg	aaaggggaat	caggtcttgt	149940
tagagatgat	tgtgagtcac	catgtggttg	ctgggatttg	aactcaggcc	ctccagaaga	150000
gcagtcagtg	ctcttaactg	ctgagctatc	tccagcccca	tttttagtat	tcttattaag	150060
tggtttccat	tttatccaaa	gatgccttta	agggcctggg	aagatggctc	agtgggcat	150120
gaacttgggt	tgtgagcatg	aagaccagag	tatcagatccc	tagcaccag	gcagatgctg	150180
aatgatgggt	gcctgcctga	gattccagga	caacggagac	agacaggggc	cctagctaac	150240
catactacac	actagctgag	ctgtgtgctc	aagagagcag	ccctggctta	ctgtgcagga	150300
tggagagtga	tcactctccac	aggcaagcac	acacctgagc	acacagacat	gcacaaagga	150360
aagaaaagtc	cttttaaggt	ggtggtgggt	ttgttgtttg	ggggtctttt	gttttgtttt	150420
tttctcccc	tccctcattg	tgatggcaca	ttcctttaat	ctcacatctg	ggacaaagag	150480
gccggaggat	ctttgtgaac	tggaggtcag	cctgttctac	atagcaagcc	catttcagcc	150540
aggacgacat	agatataccc	tgtctcaaac	agacaaaaat	tatttatttt	atatatttga	150600
atgttttgcc	tgcatgtatg	tctatgcaca	ttatgtctgg	tgcccatgaa	agccagaaga	150660
gggcatcaga	tctctcagaa	ctggaatttc	agacacttat	caagtactgc	ctgagtgcata	150720
ggaatcaaac	caaggctctc	tggaaagagca	gcaagtagtc	tttttttttt	ttaataatttt	150780
tttattacat	attttcctca	attacatttc	caatgctatc	ccaaaagtcc	cccataacct	150840
cccccccccc	ccccgagcag	caagtattct	tcattgctgg	gccatctccc	catctccttt	150900
tctagttaat	taagctgaaa	gggagggagg	tagatgttgc	ccaaacttag	gatttattga	150960
cagattaata	ctctgttagc	ctaactacac	tatagaagct	tattctttag	actttcacat	151020
tacactgtcc	agattttgcc	atcctttttg	ngtgtatatg	tctacagatc	ttaattcagc	151080
tgccaattta	tacagtgttt	ataggtattc	tttgtgacgt	ggatctttta	cccactctaa	151140
agcagtagga	tttgaaagct	gacatttatg	tggcctatgg	tcctgttaaa	tcacatttca	151200
agttagtctc	tgtggtacac	attttgggg	ctatctgcgg	ttccgcatct	cacacttttc	151260
cctctcaggg	tgtccagaag	ctgctgcaca	ctgggctgga	aggatgaagt	ggagtcagga	151320
gtgagtggaa	ttctgcagca	tcccgggtcca	gctgggagtg	aatgctgggg	tcaggaggag	151380
atgggtgaga	gggccttctc	caagggcctt	cttagtggtta	cagctctagg	caaaggcctt	151440
ctctgacaat	cttagcctgt	gcatagtttt	ttattcgaga	tgagcttgta	tgcatacact	151500
ttattggcag	taaatcagag	gttatccact	cttaggggaag	gagataggaa	tacccaaggt	151560
ggacagaggt	cattggctga	aggataacgt	actgagatgc	tcattagcac	ggggaggcat	151620
cccaggaatc	tcaggtgctt	gctgactggg	ttctctaggg	gggtgagagg	gtagcagtga	151680
tttcaccaag	gttatgtatg	gcagagggta	taggggtttc	aggctcccca	gacaaagaag	151740
gagaaggaga	agccctgctg	ataagggaag	tccccattt	tgagccactt	cagaaggcta	151800
tcaagacact	gatagacttt	gtccttaatt	agcagggccc	aacaagtgtc	tgttttcttt	151860
tctgccttca	ttggctcttt	gagccactgc	tacaaaaatat	ccaaatctgg	gccaggaagc	151920
tggctcggct	agtaaagggt	tttgctctta	agcctgaaag	cctgagtttt	cacttgattt	151980
ggtttgggtt	ttgttttttt	gagacaaggt	ttctcagcat	agccctgggt	attctggaac	152040

tactctgtga	gatcaggctc	aacttgaatt	cagagatctg	cctgcttcta	catccccgagt	152100
gcttagatta	aagttgtgcg	ccaacactgc	ccacctaaaa	aaaatatgag	gggctgggtga	152160
gatggctcag	tggttaagag	cacccgactg	ctcttccgaa	ggtccgaagt	tcaaatccca	152220
gcaaccacat	ggtggctcac	aaccacctgt	gatgagatct	gatgccctct	tctgggtgcat	152280
ctgaagacag	ctacgggtgt	acttacatat	aataataaat	aaatcttaaa	aaaaaaaaaa	152340
aagaacacta	ttcacaaggt	acaacaacag	tgtactagga	aactttaaaa	tagcccatca	152400
ttcattctag	gggaaaaatc	tttttaaaact	ttagtgtgta	agaaagagag	aggggctgta	152460
gaaaggccac	agcacgtgta	tccagggttag	gagtcaactt	ttcagaagcg	gagtctcccc	152520
ttctacctgt	ttctgaggca	gtctcttggt	tctgccctac	actttgtaca	tgaacttcaa	152580
gatggttggt	ctatctctgc	ctctcatggt	gccctatgca	tgctgagctt	actgatgcca	152640
gccaccacat	aagcatggca	ccagcatgga	gccaaaggga	ggcatctggc	tcacatggag	152700
agttacccat	caagccatct	tgctagcccc	agaaaatgta	tttttgacag	gtgtgggtggt	152760
gcacatatct	aatccagca	ctcaggaggc	agaggcaggc	agatctctgt	gtctgaagcc	152820
agcccagttt	acaaatcaag	tcccagaata	gccaaaggcta	catagagaaa	ccctgttttg	152880
aaaaacaaac	atccttctgt	gttccagtc	caatgactgc	tgtacaata	atatgaggat	152940
ttgggctgt	caaaaatcac	aagtagcaag	gtgtaatggg	caggccttta	gtctcagcac	153000
ttgggaggca	gaggcaggag	gatctctgtg	agtttagcac	agccagggt	gttacacaga	153060
gaaacctgt	ctcaaaaaaa	ccaagcaaaa	atagaattac	aagttaacca	gggtattggt	153120
gttaatatga	aatagcagaa	ctcaagatag	ctaataaac	aaaggattct	attttataaa	153180
ggagacattc	catactgaaa	tatatgcaga	gcctgatgct	tgccctagcaa	ctgaactaca	153240
cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	ngcctgggca	agagtaagca	153360
aaccgtcagg	ctcagaggca	ggatggcagt	gatgtgtttt	ggggcacaag	gcccccttct	153420
attaaagcac	caaactctgt	actaaaaaca	gccctctgct	ctcgatctgc	agagatacga	153480
gagaggcaca	catagcttcc	gggcccttgc	cctctgctct	cccggctcagt	tcctggactc	153540
ctgggagaag	cttgaagctc	aagaactggc	cgctgttgcc	tttgtgctga	cgccagaacc	153600
agcagctatt	cagcagctgc	ggctcttggg	caagcttggg	agtagctcgt	tccttctcct	153660
tctgcagggg	aaagacaagg	agatggcact	gagctgggta	gccaggatgt	gggaagtaaa	153720
ttgtgtctat	gtgggtgagg	gaagtgccgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	153780
gtgtgtgtgt	gtgtgacaga	gacagaaaag	gagagtgcac	gttcttgtgt	gttctgtgta	153840
gtcagtgtct	cctttggctc	ctcctgccaa	aaagcatgct	gtgttctcag	agtcaagctt	153900
ggccaggctc	gcccacccca	ccaaagccca	ggacttgcca	ccccattaa	aatcggggt	153960
taggaaataa	aaaacagggt	tcttgtccaa	gagaattttt	attattattt	tttctctctt	154020
aatggatttc	agcatttctt	actcctcagt	atcctctctg	gtagggaatt	caggtctgtc	154080
tatgcagagc	acaagagact	gtgcttgcca	gaaaccttgc	ggccaacagc	cctattccct	154140
gactgggctt	gcctgcaggc	tgccctcttg	gctgacccct	gagtctggcc	ctctgacctc	154200
tgcccgctct	ggggtcatcc	aggaggagca	ggaccaactg	gatacagggt	tcctgagatg	154260
gctactgaaa	cacctcatc	tgctaaggcc	actagatttt	tttccactg	ccagactact	154320
cagggccctc	agtaggtcac	tgcccaaaga	gcctagatgt	taaaactgca	gctaaagcct	154380
ctcctgaggc	cagagctcag	agcctccctg	gcctgcacaa	agtgtctaaga	aggacattgt	154440
ccatccaacc	cattggacta	acactgtaga	cgctgccttg	tctgccagct	tgaaaacccc	154500
agaaacctct	tccccagctc	cgccctagctt	gcctcagcga	ccctgacatc	cacttctctc	154560
gctaattgtc	gcagcttcta	cagtagggtt	gacgggggtg	tccgggggtg	cagacaggcg	154620
tgtgttgcat	ataaaaaacga	ggtgatgttc	taagtatcta	agaatgttgg	ttccctgaag	154680
tgattcttgc	tgcttctctt	ccttcccagc	tcttcccact	gacacctttg	ccccagcaa	154740
gcccagggac	tattgttgct	ggctgggggt	ctgaacaaaa	ttgccgcagt	tttttgtttt	154800
tttttttctt	gctagaagtt	accgatacag	tcccttaattg	agcaaataa	ggttcccgtg	154860
tagtttataa	acatgataag	acatacagtt	tggttaaggag	tggttggtgg	gacctgtgca	154920
tttatatatt	tatatatata	tatatattat	gtgggtgtgt	gggcagagtg	aggatatata	154980
taagtggaca	taggtataaa	actgcacctt	ctgtgtgact	ctcattgcga	gtacagttct	155040
aaatgtcatc	cactggcgat	ctctcctttg	gtgattgggt	cttggacccc	agcaggctctg	155100
ccgggggctg	cctgagacgt	caggaatgag	aggcattacc	cccattggata	gggactgagg	155160
gtggcatagg	gttggacagg	gcaggttaac	taagtatctt	cagacaagag	ccaagagtgc	155220
tctgagattg	ctgggtgccc	cagctggctc	tgggagagcc	ttgttctgag	tcctgctcct	155280
tccaaaacca	gcagggtcct	tcagcccttc	tctccaaatg	accaggcttc	cgagagagccc	155340
agcttcttca	aggggctcat	gtcccagcac	cactaatgac	tcactttgctg	tgccctttgac	155400
cactgtgctg	gagtggtatc	ggtccagagg	cgctcggtca	ggacagccga	gtgagacgtg	155460
atacccttcc	cgtctacggc	tgtacatttt	ggccttataa	tccaccagga	agggctcaga	155520
cggtggctga	aggcttcagc	agccttttcc	tgaacccag	cagatcttcc	acttaggaaa	155580
aaaaaaagaa	agaaagaaag	aaaagaagaa	aaaaattctg	ttccttgctg	gacatgtggc	155640
agtgtctgggt	gacggagccc	tggttctcca	gaggagagtg	actgccagcc	ccagtatcca	155700
ggccaggagt	ggggcccaa	gggctgtgcc	tgaggatgcc	tgctgcaccc	cactgggggc	155760
acggatgggg	gtcctgcgag	ctcctcttgc	gttcctcttg	ggctgtgcgg	tggtgtgat	155820
gtcaaagctg	aacttgtgct	gatagtcggg	ggcatagtcc	tgcatgtcgg	taagctcttt	155880

cccagagctc	accttagaga	tctggttccg	gttcctgtgg	ctggtgcagt	tcttgccctgc	155940
cttctttaa	cctgacctgg	agccaggcgg	atggccatgt	gggtggcctt	tgtccctgga	156000
ggccccatgg	gacggatggt	gctccttgcg	ggcagccctg	tcagaggtgg	taagcgtgtg	156060
agacttgatc	tggtgaggag	acactgggtcc	tgtgcagttc	cggaagtcct	ccaccctcag	156120
cagcttcaga	tcttgccctt	gccgcagctc	gggggtcgcg	caggggacag	cagagctaga	156180
gccacggaac	cttcgcagcc	attcccacag	ggaacgtgcc	cggcagccac	agtcccaagc	156240
attcccattg	aggcgaagga	actccaaggc	caccaggggg	gccagacagt	caccctgcag	156300
ctcagtgagg	ctgttggtga	agagaaagag	ggtggttagc	ctgtggaggt	catggaaagc	156360
cttgtggtga	acccactgta	gctggttctc	atgcagcagc	aaccggtcca	ggttcaccag	156420
gccccggaag	atgccttgge	ccagggtcca	tagcttggtt	ccatggagaa	acaagtgcag	156480
gagattgacc	aggtccacaa	agatgtcatc	ttggaggtac	tcgatatggt	tgtcctgcaa	156540
gtagagatac	tgagggtgtg	gcaggccacc	aaagatgcct	gcgggcaggg	gcctcagtc	156600
acacttatag	aggtagaggg	cgtgaagctt	caccaggcct	tggaaggtct	cgggtgccag	156660
cgttcgcagc	tgtcgggtgt	ctccaaggte	tagctcctcc	agatgcacaa	agccctcgaa	156720
ggtgttgga	gcaatgaaag	tgatgttggt	ggagtagatc	cagaggggtga	ccatggcggg	156780
gctgaagtgg	ccctgctgga	ggaaggtgat	gcgattgttc	tgcaaggaag	tgcgctcact	156840
gtcctctggg	atgcctcccg	ggatggcagc	aaagtgtgtg	gcctggcagc	tgacagtcat	156900
gggcgcaggg	tagcacacac	agtcctcagg	acaaccacca	cccagaggta	gctctccagc	156960
gagcagcaac	agcagcaatt	ccacacagca	ccctggtggg	gagagacaga	acagcagtga	157020
ggggctgccc	agaggaggtg	gagatagatg	gggaacagag	ggtggaatgg	gggactggca	157080
aatgactctg	ttgggtcaca	gagggtctgt	cctctgtatt	gtatgcaggg	gtcccttgga	157140
cagggcattt	ggggccaagg	cccacattat	ctcctcact	ctttagctct	gtccctaaag	157200
tctctaattc	catccgaaca	ctccttcaga	ctgtcagccc	caccgcaggg	tggaacacac	157260
ttgtgaacac	aggcgagtcg	gcctccggct	ctgggtccgg	ctctgccact	cgctcactgt	157320
tagctgcctt	agcaaggaat	gactctaaca	aagcaaattc	ggagtcctga	atgatcactt	157380
tatttaaata	attcctcaaa	ataaagaaag	catttagtcc	atggtacca	agcatgcctc	157440
aataagcgcc	tctttcacac	tgtggtacaa	aaacttcaaa	cctacaactc	ctccatggct	157500
gtttcctcat	gagttaaaca	cagttcacag	ggctgtgtgt	acagacaagg	cacatttctg	157560
tgaggggctg	tgagtgcac	cagggtgac	gcacgaggct	tcccttgggg	ttcacagtac	157620
tgccagatcg	aggctgcatg	ctcctctccc	ccattcacac	ccccccctcc	tgtgctggag	157680
attgccaggc	tgtggtgta	aaaccgggcc	ttgcctcttg	actgtccaga	gcatttctc	157740
tgtagcttcc	ctctaattgg	gcattaatta	ggcattcggt	aatggatcct	taaaataatt	157800
attttcggat	gtgtccagcc	tgtggtcggt	taataggcct	atgctcatta	tggaagccgc	157860
ctcattatgg	acgattgtca	ttacctgcct	ttttccaggg	tcacagctgc	cccaagtggc	157920
ccagcaggcg	cgtcaggaag	atggggacag	gctccaggcc	tacgggcgcc	caccctgaag	157980
ggccaggcag	ccacgacctc	tgtcgcctca	gttgccctct	tgcctctct	tttccagctt	158040
gttcagctgg	gactcttggt	agagccaggg	ccccggggg	aatatgagct	gagctgaatc	158100
ttcttgctgc	tagctgtgct	cagagcaagt	ggaaggagca	gggaccttct	gaccaggctt	158160
cccacttggg	gtcccaggcc	cagggtactg	ccaggccccg	gcagagtagg	ttgccacctt	158220
gacttctgac	gccccccccc	cattcccaac	agaaacagca	tcgtaagttg	acagctccca	158280
gctgttggga	gttatgggct	cccagagagt	ggcagctgct	tctcgtccct	gtaatcacc	158340
ggcttcagct	agaatgtttc	tagcacataa	aatcatcgc	atataattta	gtttttgcat	158400
aattgggttc	agttgtgatt	tcagagcaat	tatgacctca	gcagcagggg	tgccacagca	158460
taggacccct	ttctgggccc	ggccatgccc	tgcaggccct	ctgcagtgtc	ttctgcccac	158520
cggcccttag	acagcatgca	ggctataacc	atctctgcct	caatttctct	ctccaaaacg	158580
tgtctagatg	ttactctgtc	gatcttctct	ctcagcatcc	tgggtgtggc	ctccagcctg	158640
ccagcctctg	tcctagggat	cctgtgtgtc	agagggaggc	acagtcggag	ggaggggagc	158700
ctgcctgtgt	ccaccagcac	tcacactggc	tgcacagtcc	acagaccac	agctccaacc	158760
tccctgcttg	gcttgacccc	tctcttccag	gaaggccatt	cttgccagaa	cctttcccaa	158820
cgggtcccctg	ggaaagcctg	gactctaggt	tcaaggacat	tcagatgct	tgccccacat	158880
tttatgctgg	atgagacaca	gcagagcctt	cttcactggg	gggtcctgtg	aaaatgaaag	158940
cttttcttcc	ccggcctgca	gctgcaggca	ggttaggggt	gcagtgggct	tatcactaat	159000
accattcgac	atttgtacag	ctcatcgagg	tttacagagg	gcttttgttg	taccctcaac	159060
ttcccctgtg	ttctcccacc	tactgtggct	gctctgtctc	tgtgcacccc	aaaagaatct	159120
ggaagtccct	tggggagatt	acccccctta	catagggggc	tccaaggaat	acaggctaca	159180
gctctattta	ggaaaaaaa	aatcaaaact	aaccaaact	caggtgtgga	cttagtaacc	159240
agtttataaa	cataccgtgg	caagtggagg	aggcaggcgg	cagaacggca	tgaaggtagg	159300
actgggggtt	tcctctaaa	agggcaatgg	ggggacaaag	ggactctagc	caagacttag	159360
tgcaagaacca	ggctcagttc	ccctgttatc	tcaaggctat	catcactagg	aggcctaagg	159420
caatggacca	caaggacctt	gtcctttag	gacagtcact	tcctgcagtg	aagtgtctct	159480
ctggaagcat	actaatagga	tgtaggctca	ggacagctgg	tctctgtcct	ttagtatttt	159540
tccacatgcc	agggatgtta	accttccaag	cctccatctc	ttctaattggg	gggggggtgt	159600
tgaggggctc	agccactctg	catccatgtc	tttgaaagcc	agtgggtatta	ctccaggacc	159660
ctgagcaagc	tgtcctagtc	agccctggcc	acttctggac	tccttgcttg	agtcagtagg	159720

tgccaatcct	aggattgtca	ccagcagggt	tcttcctagg	gaggcaagca	ctgtatcacc	159780
atggcgccct	ctatgcccc	tctatgaggc	ccttggggagc	cccgccccac	tgattgacctg	159840
attaatgtac	caacaatgag	gatggagcct	ttgccatgca	ttttaacatt	gcaaattagc	159900
aggaatccaa	gtctctgtgg	agggggccctg	cacctcttct	gccagactca	tcaagcgcct	159960
cttgggcagg	gctgccttct	acttgagggg	gcggaaggga	gaagaccag	ttccactctc	160020
cttccccctcc	aggaggtgcc	cttcatcgctg	ttctgcttcg	ttactctcaa	gcctccggcc	160080
tcccacgcac	gtgagctccc	aaggggctct	agacagtgtg	ggaagggtct	ttccattcat	160140
acttgcccc	tagtctagga	gagccattgga	acagcctccg	tcattccttc	gacaatgagc	160200
atcatgcccc	atattgcatat	gcggtggcaa	taccctgggtg	ggtaccagga	gagtataggg	160260
gaaattaaga	gaggggccta	aggaaagcct	ctgctatccc	tgggctacca	gtcagcattg	160320
cttggtcact	gatccccctct	gtaacaccag	cccttctgca	acctgccaga	gtttttgacc	160380
tttgaactag	ggctgagaag	ggtctgctct	gttcagctgc	ccttggtggg	aggggaatct	160440
gctcagacct	cagcacacac	tcaacagaag	gcatgcaagc	aaggagcta	gcagtggcct	160500
tgggtcagct	ggcaagcccc	aaactcttcc	tgccaagctg	agcatgaaaa	gccacctcac	160560
catggtccca	tgggaccaga	cctggttagga	taggtggcaa	ggctaaggca	gcggaatagc	160620
atgtgcaaa	gcaactgggt	gggaaagggc	ctgtgcttct	cacccccctct	aatgggtgcag	160680
agcctccaag	gaatactgta	acctcagctc	agctgggctc	gggtggccag	agagcttggc	160740
accagaacca	gcatcaacag	ggcctgtctg	ctaaaccag	acctcacaag	ccagtttagt	160800
aggggccctg	tagcaccctg	gccaccagaa	ctaaccagga	agatctgacg	ctgggaatat	160860
gtctttaaat	aaaagccctt	ccggaagcca	catttgacaca	gaagaaaatg	aggtgccag	160920
agcatcagtg	ggctggttgc	agctggagaa	cacagcaggg	ggacaggctc	taccaagcta	160980
ccctgccttc	aggctggggc	tctagccagc	tccctgatgc	ctggagtagg	taaagcagcc	161040
tcgaaatggg	ctgggtcagc	tttttcaggc	tccaaagggt	caggacagct	gtgcagctt	161100
agcacccaag	ggggctgccc	ctctaccctt	aagtagaggc	atccccatgg	cccttgggca	161160
ggctcagtg	tctctctgaa	gctttgtagg	ctgcttcttg	gccatgtagc	caatctctct	161220
ggccttcagt	cctccctccc	tgcccccagc	ctggccagct	gctcttctct	gagcaatcga	161280
tgtaaacga	atgctctctt	gctgtgggga	tggcgccctc	aggccaggcc	agctgcactc	161340
ctggggctgc	tggcgccctca	ggccacttgg	cacttgtgcc	acttgtgttc	taaacacagg	161400
ctccttccctg	gctcgccctt	gaagacaaga	aagctggcca	gggaacagct	gggctcccat	161460
ctcagcctcc	actgctgtgc	agagcggccg	gcagcctcct	atccatggct	gtgagtagaa	161520
cagggtctgtg	gagccagagg	cctaagttga	atcctggctg	ctccttttaa	tgcttgcagg	161580
agcctcgttt	tctcactctg	caaaatgggg	cagtcattgg	aagctcagcc	agtcctccag	161640
cccacagata	ctggtaccca	ctgccttcc	ccacatctcc	atctgtctaa	ctgaaaccat	161700
ccaaaccaag	ctcttctctt	cctctccggc	ctcctccgtt	cacaacttct	ccatcttggg	161760
taaagatgg	ccctttacat	cagctgcctg	ggaggaacac	ctcagaatca	ccgtggctca	161820
ctctggggaa	attctgctgg	ctagatttta	gaatgtatcc	agtatctatc	cacttatcat	161880
actcgctatt	gctaccattc	acccagtagc	ctcctggatg	ccctcccccc	ccccccgcaa	161940
gccctgcctc	ctcaccctca	cacctcctc	aacaggaact	agggtagtcc	agggaaagtg	162000
agtcaggaag	ggctgctcct	tagtctgcat	cctccagagt	gcccattctaa	ctagaagctg	162060
ccccagggtc	ttctccagc	ccagaggccc	tgcctctccc	tggctacaca	gaccttgctg	162120
ccgtccctcc	cacatgccag	ccctcagcct	ccctcctgcc	tttgctccatg	ctgttccatc	162180
tacctggacc	ggttcccgat	gtgtctgcag	ggctgcttcc	cagagaagcc	actcttgagc	162240
agcgatttgc	aaggagcttc	tttccctggga	cattttactc	tacagcacgc	gacctcttgg	162300
acagcacgac	ataagtcaact	tgtttccctt	atgtccgctg	ccactttgtt	ctgatgagtg	162360
tgctccctgc	acacagtagg	tgctcattaa	cagttctggg	ggaggggaatt	accttctcaa	162420
gtctctggag	aattgaatga	tgacactcag	gaagccctag	gctcaagcct	ggggcctgga	162480
tcatagtagg	tgctcgataa	atgttggttg	taattagtcc	tgggagactc	agagccttca	162540
ggagaacaga	cacctgaact	tggctcacat	caagactcct	taggcccata	aaggaagtga	162600
ccgtttgttg	gatgagcact	ctgagaagga	cccaatacca	gtcctttcct	gggcaagggg	162660
aaatagactg	ggactgggga	gtttcccaag	gtatgggatt	ttcagcacta	aactggaata	162720
atgctaaaga	aaaaaaaaag	ttagtcactc	taaaaggggc	caggaccaaa	acctttcaaa	162780
cagaaatgtc	tgggtttatg	aagagaggaa	gccagatatg	gtggggcaca	tctttaatcc	162840
aggtgcttgg	gaggcaaaga	cagatggagc	tctgtgagtt	tgaggccagc	ctattctaca	162900
aagtgaattc	taggacagcc	aaggctacat	agagaaaccc	acttgacttg	ccacccaaat	162960
taaaaatctt	agtgggggag	cagtcaagga	ggaacacac	acacacacac	acacacacac	163020
acacacacac	acgttagtat	aatatcatac	tatggtctctg	tgcttgagct	ccaggaatga	163080
gggctgaact	cagagtgtta	gtgtgtgcta	gtggatat	gagctctgta	tttatgtgca	163140
tgtctgtgta	gatgtgtacc	tgaggtgttt	atgtgtacac	aggtgttggc	ctgttgcata	163200
tggatggaga	catggttgtg	tttgctaggc	atttgcagtg	gtctgagttc	atgcacataa	163260
actcacatct	acctctggag	actgagagtg	acaaccagg	gcccttttat	cctgcagcac	163320
cccaggccca	gcaccccgac	ccagcatccc	aggccagca	ctccagaccc	agcaccacag	163380
gcccagcatc	ccacgcccag	catcccaggc	ccagcacc	agacccagca	tcccaggccc	163440
agcaccacag	acccagcacc	ccaggcccag	catcccaggc	ccagcacc	gaccagcat	163500
cccaggccca	gcaccccgag	tccaagcact	caatgccag	caccccgacc	cagcatccca	163560

ggccccagcac	cccaggccca	gcatccaagg	cccagcatcc	cagggacagc	accccaggcc	163620
cagcatccca	ggccccagcat	cccagggaca	gcaccccagg	cccagcatcc	caggtccagc	163680
atcccaggga	cagcacccca	ggccccagtat	cccagggaca	gcaccccagg	cccagtatcc	163740
cagggacagc	accccaggcc	cagtatccca	gggacagcac	cccaggcaca	gtatcccaca	163800
tggaggcagc	acatactgaa	gatagggaat	gtctctgagg	cctcttatct	tggtccttac	163860
cctcattgct	ttcagcacct	gtctctctca	cactcggaat	caaacaccct	gtgcaggttc	163920
tcccagtacc	aggattcccc	tcagctgagg	aatgggtagc	taccattttg	gcttttgtct	163980
gtctgggggt	ggcagcccca	tgctaattgg	actgacagtt	tctcctgaga	gcaatttggg	164040
cagcacatcc	tgcccattag	gcctaacctt	gcctgcaggg	gtgtgctgtg	ggggcaggga	164100
tggagcctac	cctgtatagc	tctgtattga	ggcactcccc	caagctatga	cccatgccag	164160
tgggagtcac	ttcacctagg	caactccaga	tgggcacaaa	aatctctcca	ataagggtag	164220
gtatgggaat	aggtaaggag	agcatagtga	gcctggctgg	gcacctgaga	cctgagcagc	164280
ctgcacggga	gattgtgtca	ctgtgggtcc	agactgccaa	gacatcttgg	ctttcacccc	164340
aactcaggat	ggtccagaat	ccagagctct	taagagagca	gatgctgaga	ggcacttaac	164400
ccagggctaa	gaccttccct	tggacgggtt	cttggctttc	tactctgtcc	tctgtcccag	164460
tctgtcatcc	ccatctgtgc	ctaacagctc	tctgtggaaa	acatgaggcg	tatgagctct	164520
ctacttctcc	cagcatccca	tgcccgcacc	ccagctcact	gtgtgcccct	atgttactca	164580
aatcttctgc	taggtttgag	ggccccagg	tgaggctgtg	ggtttccctc	catctgtccc	164640
tcccttttac	caccaccact	aatectcttc	ctcctcttcc	tcctcttcc	cctcctcctc	164700
ttctctnctn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164820
aaattgaact	tacagaggaa	acacaaatga	gctgtaatga	cagaaaagga	atactatttc	164880
ccttgtgatc	agaaaagctat	agattaaacc	ccagcccttt	ttgcctttgt	ttttcatatt	164940
tagtggttct	ggggactgaa	cccaagggaag	gacagctgct	tggaaatgtg	gagacctctc	165000
aggcattgct	tatgctaagc	tgacaaaggc	tttctgaatg	catatggtag	tatctattca	165060
tcttaagatg	cctataacca	tgggtggagc	aattctctaa	gagtctgtct	gaacttctta	165120
agcatgtgct	caaaaacaca	cataaagatt	gtttactgaa	gcagtatttg	caacagtga	165180
acaaaaacca	acaacaacaa	aaagtactta	atagctcatt	agtaggaaaa	agaacaacaa	165240
attgggacac	atcttatacta	ttctcgttct	atcacacgac	tattaagaag	attaagggtg	165300
gaggatgaaa	agccccctgca	gaacaacgca	cacaatgaaa	ttgaatagag	aataaagaaa	165360
aaaatatata	gtgtattttat	gtaaaacata	cataaacaca	taaaaagaat	agaaaactct	165420
ttgactgtat	atctttgaaa	agaaccagca	gagtgcattg	gtttataaat	gaaaagttcc	165480
cccaaaattc	ctgagttgga	ggttcaatgc	agagttcata	ctcttaggaa	gtaactaaat	165540
caggaatgct	ctgatcta	caatggattg	accaactgat	agattaataa	tctgaaggca	165600
acacttccag	gaggcagaaa	caagggtgagg	cctaactgga	ggaagtttgt	caccagaggc	165660
atgtccttgg	agaggctacc	tcactccttg	tccttccctt	tctccgcttc	ctggctgtaa	165720
aagtattttat	ttgactgatg	taactgtcat	cttggaggct	tactggctcc	atcagctaac	165780
ctaggcctag	ccctggaagc	ttctagcttc	catacaatct	aatccaagcc	tagaatgttc	165840
cagcctttag	gacttgctgc	tgagatcacc	gtttcctgtt	ctttctgaac	tctagctggc	165900
tgattcagtc	cccctgttcc	gggctcaaac	tcctctcccc	gatgatttta	ttcacaatct	165960
gtcttttctc	ttggcctctg	aattgctctg	cttgggtctca	aactaactct	agcaatcttt	166020
tctaatctct	tgtctccttc	acactctctt	gcttggtctg	gtctttactgt	gtctagtttg	166080
ttctctcttc	catccttctc	tgtaaaagctc	tcocggtaaa	cctgcctcct	cctccccctc	166140
tgtgcgcttc	tactctccct	ctcagctcta	ctgcactgct	ctccaagct	ctcctgtatc	166200
ctgtgctgca	ctctcttctc	cggtaccacc	tgtgtctccc	ttacgtagct	tccttttcc	166260
ctctcttctc	ctgagggttg	ggcagatcct	atcctgtcaa	acctttctct	gattcttcc	166320
tttgctctgc	actcaattag	acatcacttt	caagcaggag	tgccctcctc	acaaaccaac	166380
tttaccttca	ttgtttcaaa	ttaaagggtga	gtactaaagg	tgtgtctctt	tttcagccag	166440
tgagagtaaa	gatgtgtgct	aataaggctg	agccaactct	agctagaaat	agtttctttt	166500
tctccataaa	taacagaatc	ttagggttca	caatacgatc	aaatatcctg	agacagctgg	166560
ctgccatgtg	gtgaagagaa	tgacagagat	taaggatgtg	ctcttgaggt	ttcagatggg	166620
aataagcatt	ttcttgggag	ttgagactgt	gtcattcctg	tgacactgtg	acaaaggact	166680
cgactacatt	ttgccatgcc	ttgagactgt	ggagggtgga	gcgatggatt	aatacactgg	166740
agtgaatttt	aaggcagcca	aggctgtggg	tggactgtta	ctagctacgt	ttagctggat	166800
ttatatatta	aattgggaac	aaaaaagcag	agtagaaagg	acagtttttg	cagaaggagg	166860
tttcttcata	tacaacttag	tttttatagt	tagcttttta	tgttgctatg	acaaaatac	166920
cttatggaaa	ctgaagaata	aaaattttta	ctgaactctt	cttcacccca	gaacccgacc	166980
cctcccatct	agagattggt	cccggaaacac	tcctgaactc	ttcaccaccag	aatgctttcc	167040
tgaactcctc	accctagagt	tcgaaccctc	ccaactaaaa	actgttccaa	gaacattttt	167100
gagataaggg	cctcctaaaa	caacctcaaa	atgaaccggg	tacattgcca	aataatagga	167160
catgacccct	tagttacgta	gattcccttg	gcagaacccc	ttgtcccttg	acagaacccc	167220
ctagtgatgt	aaacttgtac	tttccctgcc	cagctctccc	cccttgagtt	ttactatata	167280
agcctatgaa	aaatttgggt	ggctcgtcgt	tctcctctac	accactaggt	gcatgagttt	167340
cgacccccaga	gctctgggtct	atgttccatg	tgctttcttg	ctgttgttct	attaaatctt	167400

gccttctaca	ttttgagtac	ggtctcagtg	tcttcttggg	tccgcggctg	tcccggggct	167460
tgagtgcctt	agtgagggtc	tcccttcggg	ggtctttcat	tttgggtgcat	tgcccgggaa	167520
acagcgcgac	cacccagagg	tcctagaccc	acttagaggt	aaggttcttt	gttctgtttt	167580
ggtctgatgt	ttgtgttctg	tttctaagtt	tggtgcgac	gcagtttcgg	ttttgcggat	167640
gctcagtgag	accgcgctcc	gagagggaa	gcgggggtga	taaggataga	cgtgtccagg	167700
tgteccaccg	cogttcacc	tgggagacgt	cccaggaaaa	acaggggagg	accaggagcg	167760
cctgggtggac	ccctttggag	gccaagagac	catttgggg	tgcgagatcg	tgggtttgag	167820
tcccacctcg	tgcccagttg	cgagatcgtg	ggttcgagtc	ccacctcgcg	ttttgttgcg	167880
agaccgtggg	ttcaagtccc	acctcgcggt	tggtcacgag	atcgtgggtt	cgagtcccac	167940
ctcgtgcaga	gggtctcaat	cgcccgccct	tagaaaaggcc	atctgattct	ttgagttgct	168000
tgtggtcgac	gcagagtcgc	cgccgtttct	ggtttctttt	ttgtcttagt	ctcgtgtccg	168060
ctcttgttgt	gtctactgtt	tttctagaaa	tgggacaatc	tgtgtccact	ccctttcttc	168120
tgactctgga	gcattggaag	gaggtgcggg	tcagagccca	caaccagtcg	gtggaagtca	168180
gaaaggggtc	gtggcagacc	ttttgcgcct	ccgagtgggc	aacgttttaga	gtaggctggc	168240
cacctgaggg	tgcttttgac	ttgtcactaa	tcgctgccgt	caggcgaatt	gtttttcagg	168300
aggaaggggg	tcaccctgat	cagatccccc	acattgtgac	ctggcagaat	ctcgtccaat	168360
tcccacctcc	gtgggtcaag	ccttggaccc	caaactcttc	gaaactgacg	gtcgcgggtg	168420
cccagtcctga	tgacgcccga	aagtctggcc	catcagcacc	cccccaagatc	tatccagaga	168480
ttgacgacct	cctctggata	gactcccaac	ctccccctta	ccccctgccc	caacagccac	168540
ctgcagctgc	cccaccacag	ggaccaatag	cgagaggggc	tcaggggaccg	gcggggggaga	168600
ctcggagtcg	ccgagggccga	agccccgggg	aggaaggggg	gccagactca	acagttgcct	168660
tgccactcag	agcacatgtg	agagggccag	caccagacc	taatgatctc	attctcttac	168720
agtaactggcc	tttttctctc	tctgatttat	ataattgaaa	aactaaccac	cctcccttct	168780
cagagaaccc	ctctggactt	actgggctcc	ttgagtcact	tatgttctcc	catcaaccac	168840
cttgggatga	ttgtcagcag	cttttgagg	ttctttttac	aacagaagaa	agaaaaagaa	168900
tcctcataga	ggcgagaaaa	aatgttcttg	gagaggaagg	cacacccact	gccctcccta	168960
acctcgtgga	cgaggctttc	ccctgaacc	gccccaaactg	ggactacaac	accgcggaag	169020
gtaggggacg	cctcctgttc	tatcgccgga	ctctagtggc	aggtctcaga	ggagccgcta	169080
gacggcccac	caatttggct	aaggtaaagag	aggtcttgca	ggggcagact	gaaccaccct	169140
cagtcttctc	tgagcgtcta	atggaggcat	ataggagata	cacccttttt	gaccctttgt	169200
cagaggggca	gagagccgct	gtagccatgg	ccttcattgg	tcagtccgct	cccgacatta	169260
agaaaaagct	gcaaaggctg	gaggggctcc	aagatcatat	gctccaagat	ttagtaaaag	169320
aagcagaaaa	agtctatcat	aagagggaaa	cagaagaaga	gaggcaggag	agagagaaga	169380
aagaaataga	ggagagggaa	aatagacggg	atcgccgtca	ggagagaaat	ctgagtaaaa	169440
ttttggccgc	agttgtgaat	gatagacagt	caggaaaagg	taaaataggg	ctcctgggca	169500
acagggcagt	gaaaccgcaa	ggtggcaaaa	agataccact	ggaaaaagac	caatgcgcct	169560
attgcaaaga	gaaaggacac	tgggctagag	attgcccata	aaagcgggag	cgatccaagg	169620
tcctaaccct	agaagatgat	tagggaagtc	ggggctcaga	ccccctccct	gagcctaggg	169680
taactttgtc	cgtggagggg	actcccgtca	acttctctgat	agacaccgga	gcagaacatt	169740
cagtactcac	taacccctta	ggcaagctag	gctccaaaaa	gaccatggta	attggagcca	169800
ctggtagtaa	atttttaccc	tggaagacca	aacgagctct	tcagatagac	aaaaatatag	169860
tgacccactc	ctttctggtg	atacctgagt	gcctgtctcc	cctcttgggg	cgcatcttgc	169920
taaccaaact	aaaggctcaa	gtccaattta	cttcagaagg	cccacaagta	agctggggaa	169980
aggccctgtg	tgccctgcctt	gtcctcaaca	cagaaaaaga	gtaccgggtg	catgaagaac	170040
aacccaaaaa	tgagctctct	tcaggttggc	taactgcgtt	ccccaatgtc	tgggcagaac	170100
aagcaggaat	ggggttggct	aaacaagtgc	ctccggttgt	ggtagaactt	aaagctgatg	170160
ccacccccat	ttcggtaaaa	caatacccca	tgagcaagga	agctagaaaa	ggcatccggc	170220
ctcatatcca	gaggttgctg	ggccaaggag	ttttagtggc	ctgtcagtc	ccctggaata	170280
caccacttct	gccggttcaa	aaaccaggga	ccaatgacta	tcgcccggta	caagacctcc	170340
gggaggttaa	caaaagggtc	ctggacattc	acccacagct	cccgaaccgg	tacaatttat	170400
taagctctct	cccacctgag	agaacatggt	atacagtcct	agacttaaaa	gatgccttct	170460
tttgctctcg	tttgcacctt	aagagtcagc	tcctgtttgc	ttttaaatgg	agggacccag	170520
agggcggaca	gactgggtcaa	ctaacttggg	ctaggctacc	acaggggttc	aaaaattccc	170580
ccacctgtt	tgacgaggcc	ctccatcggt	atcttgcgcc	ttttcgcgct	cgaaaccctc	170640
agcttaccc	actacagtat	gtagatgatc	tcttggctgc	ggcggcctcg	aaggagctgt	170700
gtcaccagg	aactgagagg	ctcctcacag	aactgagtga	cttggggtat	cgagtttcgg	170760
ctaaaaaggc	acaaatctgt	caaactgagg	taaccttctt	ggggtatacc	ctccgagggg	170820
gcaaaagatg	ggtcacagag	gcccggaaaa	agactgttat	gatgatccca	tcgccaacta	170880
ccccacggca	ggtagctgag	tttctgggga	ctgctggctt	ttgtagactc	tggtatccag	170940
gctttgcaac	cctagcagca	cctctatatc	ctttgactaa	ggaaggggtt	cctttcaagt	171000
ggaaagaaga	acaccaaaga	gcttttgagg	ctatcaagtc	gtctctaagt	actgccccca	171060
cgctagcatt	accagacttg	actaagcctt	tcgctctata	tgtggacgag	agagcgggtg	171120
tagccagggg	agtattgaca	caagcactgg	gacctgaaa	aagacctgta	gcctatttgt	171180
caaaaaaatt	agatcctgtt	gctagtggat	ggccacatg	tctgaaagct	attgcagcag	171240

tagccctgct	gatcaaaagat	gctgacaaac	tgacaatggg	acagcaggtg	accgttgtag	171300
cccctcatgc	cttagaaaagt	atcgtgcgac	agccacctga	cagataagat	gacaaatgcc	171360
cgaatgacac	actatcagag	cctgctgcta	aatgagcgtg	taacctttgc	gccccctgcc	171420
atcctcaacc	cagctaccct	tctccctcta	acaaatgatt	ccgtcccgat	acatcaatgt	171480
atggacatcc	tcgctgaaga	aactgggacc	agaagtgacc	tgactgacca	accctggcct	171540
agagctccca	gttggtacac	ggacggcagc	agtttccctga	tagaggggaa	gcaaaaggct	171600
ggagctgcgg	tggtagacgg	gaaaaaggta	at ttggggcaa	gcgctttgcc	tgaagggaaca	171660
tcggcacaaa	aggctgaact	tatagcgctt	atacaagccc	tccgagaggc	taaaggtaag	171720
atcggttaata	tctacactga	cagccgatat	gcttttgcta	ccgcacacat	ccatggggcc	171780
atctacaggc	agcgagggtc	attgacctcg	gctggtaaag	acattaaaaa	caaagaaaaa	171840
attctggccc	tgttagaagc	catacatgca	cctaaaaagg	tagccatcat	ccactgcccc	171900
ggccacccaa	aaaggagaaa	acttggtggc	caagggcaac	cgaatggcag	acttagtggtc	171960
aaaaacaagt	gctcaagggt	ccatgatctt	aactgaaaaa	ggtgatccgc	ccaaaagccc	172020
tgaggatggg	aggtataaca	taaaagagct	atggtagacc	agtgatcccc	tcccatactt	172080
tttttgaaag	aaaaatagaa	ttaactcccc	aagaaggaaat	aaaatttgta	aaaggactac	172140
accaattcac	ccacctggga	gttgaaaaaa	tgatgagact	aattaaaaat	tcccgatacc	172200
aagtccccaa	cctgaagtca	gtggctcaaa	agattataga	ctcctgcaaa	ccatgtgcat	172260
tcactaatgc	aactaaagcc	tacagagaac	ctggaaagag	acaacgggga	gaccatcctg	172320
gagtgatttg	ggaggttagac	tttactgaag	ttaaactctga	aatgtatggt	aacaagtatc	172380
tgttagtatt	tgtagacacc	ttttcaggat	gggttgaggc	atttcccact	aaaacagaga	172440
ctgcccagat	tgtggccaag	aagatccttg	aagaaatcct	gccaagattt	gaaatcccta	172500
aggtaatcgg	gtccgacaat	ggaccagcct	ttgtgtccca	ggtaagtcag	ggcttgggcca	172560
ctcagtttgg	catcgattgg	aaattacact	gtgcttaccg	ccctcaaagc	tcaggacagg	172620
tagagaagat	aaataggacc	ttaaaagaga	ccttgactaa	attagccatt	gagaccggca	172680
gaaaagactg	ggtggctctc	cttcctcttg	cgctcaaaaca	cccctggctg	tttcgggctc	172740
actccttttg	aagttctgta	tggaggacct	cccccttaa	tggaagctgg	tggaacatta	172800
gtttccgact	ctgaccctgt	cttaccctcc	cttttgctta	ttcattttaa	ggccctaaaa	172860
gtgattagga	cccagatttg	ggaccaactg	aaagcagcct	ataccccagg	gaccaccgca	172920
gtacccccag	ggttccgagt	tggagacaaa	gtcttggtca	gacggcatcg	aaccggtagc	172980
cttgagccac	gggtggaagg	accctatctg	gtgttactga	caacccctac	tgcggtaaaa	173040
gttgacggaa	tcgcctcctg	gatccacgcc	tcccacgtca	agagggccgc	cagtcaagat	173100
gaagaaaaacc	acgacgacaa	ttggacagtg	gcagtcactg	acaatcctct	taagcttcgt	173160
ctgcgccgca	ggcgccactc	tagacctagg	gaaccttaac	cctcatgctc	caattcaaca	173220
gtcctgggag	gtgcttaatg	aaaaggaaaa	cattgtatgg	gcaaccactg	cagtccatcc	173280
cctctggatt	tgggtggcctg	atctcacgcc	tgacatctgt	aagttagcgg	caggatcccc	173340
caattgggac	ctctcagatc	atactgatct	tagcaaccca	ccccctgagg	agcgggtgtg	173400
cccaaattggg	atagggagca	catatgggtg	ttcggggcag	ttctaccgag	ctaactcttag	173460
agctgcacat	ttttatgttt	gccctgggtca	gttcagagagc	aaaaggcttc	aacaaaaatg	173520
cgggggggca	tcagattact	tttgtggtaa	atggacatgt	gaaacgacag	gagatgctta	173580
ctggaagccc	tcctctaaat	gggacctaat	cacggtaaaa	cgaggtagtg	gctatgataa	173640
gtcaaacgaa	ggagaaagaa	acccctataa	atatcaagag	agtgggtgcg	cttttaaaaa	173700
cagagcaccc	tcaggaccat	gcaaaagataa	atactgtaac	cccctacgta	taaggttcac	173760
cgagaacgga	aaacaacacc	gtctaagttg	gcttaaaagga	aatagggtggg	gttggcgagt	173820
atacattcca	ctaagagatc	ctgggttcat	tttcacgatc	agattgacag	tgagagaccc	173880
ggcagtgaca	ctcgtagggc	ccaacaaggt	ccttataaaa	caggggcccc	ccagtcgtac	173940
tggctcccc	aaaggtcccc	actgtaccag	ctccaccaac	tccacagccc	aacacagtgg	174000
taccctccct	aggaactaat	actctcctca	taaagcctac	cttggcttcc	ccaccgcccc	174060
taggaacaga	ggaccgtctg	gtcagtcctag	tccaaggagc	tttttttagtt	ctaaatagaa	174120
ctaaccctaa	tatgactcaa	tcagtctggt	tatgctatgc	ctctagcccc	ccttattata	174180
aaggaatagc	tcagatcagg	acttataata	ctacttcaga	tcattctcaa	tgcccttggg	174240
gaaaaaacag	aaagttgact	ctagcagcag	tttcaggaag	agggctttgt	ctgggcccggg	174300
tacctcagga	taaaaggcac	ctctgtaatc	agaccagaa	catccagtct	agcaaaaagcg	174360
gtcagtatct	ggtgcctccc	ctagacacag	tgtgggcttg	caataccggt	ctcactcctt	174420
gtgtgtctat	gtctgttttt	aatagttcca	aagatttctg	catttttggtt	cagcttatct	174480
ccagactctt	gtatcatgat	aatagttctt	tttttagataa	atttgaacat	cgggtccgct	174540
gaaaaagaga	accggttacc	tttaactttg	cagttctatt	aggattggga	gtagcagctg	174600
gagtaggtac	aggaaccgct	gccttaatta	agaccccccc	aatactatga	agaactacgt	174660
gcagttatcg	atattgatct	tagaactata	gaacagctta	taaccaaatt	agaagaatct	174720
ttacttccc	tgtccgaagt	ggtgctgcaa	aatagaaggg	aattagactt	attattcctt	174780
aaaaaaagag	gactctgtgc	tgccttaaaa	gaagaatggt	gtttttatgt	tgaccattca	174840
ggagtaatca	aagattctat	ggctaaactt	agagaacgcc	tagatatacg	taaaagagaa	174900
agaaaaagcc	aacaaagatg	gtttgaaagc	tggtttaata	agtcctcctg	gtccaccact	174960
ctcctctcca	ctatagcagg	acctttaatt	acacttatgc	ttttgcttac	ttttgggccc	175020
tgcatcctta	ataagttagt	agctttttatt	agaaaaagga	taaacgcagt	ccagggttatg	175080

gtactaaggc	aacaatatcg	ggtccttcag	gaggttgaaa	actcgctcta	agattagagc	175140
tatctcctaa	aagaagtggg	gaatgaagaa	taaaaatfff	tactgaactc	ttcttcaccc	175200
cagaacccga	cccctcccat	ctagagattg	ttcccggaa	actcctgaac	tcttcacccc	175260
agaatgcatt	cctgaactcc	tcaccctaga	gttcgaaccc	tcccaactaa	aaactgttcc	175320
tagaacattt	ttgagataag	ggcctcctaa	aacaaccgca	aaatgaaccg	ggtacattgc	175380
caataaatag	gacatgaccc	cttagttacg	tctccctccg	tggcagaacc	ccttgtcccc	175440
tgacagaacc	ccctagtgat	gtaaacttgt	actttccctg	cccagctctc	cccccttgag	175500
ttttactata	taagcctgta	aaaaatfttg	ctggctcgctg	attctcctct	acaccactag	175560
gtgcatgagt	ttcgacccca	gagctctggg	ctatgttcca	tgtgctttct	tgctgttgtt	175620
ctattaaatc	ttgccttcta	catftttgagt	acggctctcag	tgtcttcttg	ggtcgcgggc	175680
tgtcccgggg	cttgagtgc	tgagtgggg	tctccctccg	ggggctcttc	aaaactactt	175740
cagaggaaaa	atgtattctg	cctcatgggt	tcagggggtt	tccctcagca	aattcaggga	175800
agacaagatg	gaacagctca	acctgctggc	aggagggtgt	gggaaaggac	aagtgttcat	175860
tgtgtgggtg	acaggaaaca	gagagctgcc	tacagtctta	caggcctacc	accactgacc	175920
tacctctgtc	cgtcaggccc	tacatcttaa	aggatctaca	gtttattaaa	agaacactac	175980
cagataggaa	ccaagtatca	aaccaccagt	ttgtagggga	taaaaataca	aggaacacat	176040
ctcaatagga	gtgtgttcca	ggatgtggg	aaggagaaca	cagttgttta	aaagcttaac	176100
gctggccagg	agagctgcac	acctttaatt	ccatcactcg	taagagggaa	gcagggttcat	176160
ctctgtgagt	tcaaggcaag	cctgggctat	acaattctag	attagccaga	gctacatcgt	176220
aggagcctgt	ttcaaaaacaa	acaaaaccaa	accataaaaa	agcattttctg	aggttttggg	176280
tttaatcccc	atgacctcaa	atagccaaac	agctctcctc	agtccaaacc	aaactgcaaa	176340
attggagcta	gtgagatggc	tcaacatagt	aaagtctctc	ccaaaaatat	tgacaactgt	176400
agcttatctc	tggggacaca	cataatggga	gaggaccaat	ttctacaagt	tacctcttga	176460
cctccacaca	tatgcctccc	acaaataaga	aaatatatat	aataaaaaaga	aagaagtcta	176520
cagctgcaca	tggtcatgca	tgccctataat	ccagcactcc	agaggctgag	gcaggaggat	176580
tattagtttg	agatcgcata	gcaagcagta	ggctagacag	ggctacatag	tgtaaacctg	176640
ccttaaaaca	caaaaatcaa	ttaagcaaca	ataacagtaa	caaccacaac	aaaaacccaa	176700
aagagtactt	tgtagttaag	acaataccaa	aatgtttcct	ttaaggacag	ttctggaatc	176760
agcaatagcc	ttccgagtgc	tcagggatgt	ataaataact	agaaaacttc	ccctggagaa	176820
atgagcacca	gggtacactg	ctctcagagc	tgcccagaaa	gttgtttatc	ctggattcat	176880
ttcagccttc	ctaactgctc	aggcattcag	aggctcactc	tgtagttagc	aatgtctaaa	176940
aaggctaaac	tactgctcag	catggctgtg	gtactttggc	ttatcatttt	gtgactgggt	177000
ttgtagtatt	gcagaattca	agagtatatg	catcatgaaa	gtttccacca	agttcctgat	177060
ccagtcacct	cttaaaagggt	ggatgcacca	agtgcctttg	gggtgataaa	ttatatccaa	177120
ataatggtat	tccaccctaa	tccccaaaaga	cttctggcca	tctcataatg	taaaaatgctg	177180
agccatcgca	ccagcccatg	gccttgaact	cttgatgggtc	ctgtctcagc	ctgtgtttgg	177240
attataaaatc	tgttggttgag	gtattccttt	gctgataata	caagcaaatt	cttcaagctt	177300
ccatcctaga	ctgaagacca	gcagctctcc	aggagtcctc	aatgcagact	ggcccagctg	177360
ggacattgag	cctcatggac	tcagccgcta	ctagattcgc	aacctattca	gacaagccac	177420
tgttggaacta	cccagacaat	actatgtaag	ccaatcccat	tttaatacac	atattcatct	177480
gggtgtgtgg	cacacacctc	tactcccagc	acgcaagagg	cagaggcagg	cagatctctg	177540
atfttcgaggc	ctggtctata	gagtgaattc	caggccagcc	agggctacac	agagaaaaacc	177600
tgtttcaaca	aaaccaaacc	cgtaaatcca	ttctatcagc	tctatttcct	tagagaattc	177660
taatacatgt	gggtaccagg	ggttgaactc	aaagtcttca	tgtttacgta	gcaagtttcc	177720
ttctgctagc	ctagtgaagc	tgaggcagggt	acagccgggt	ctttactgct	ccttgcaaat	177780
ggtcctcctg	agctttcctt	tgagagccta	caaagaactc	tttttttctt	taggtctcca	177840
ggtttttggtc	ttaagagggt	ctggacttggt	atctgtagct	gtcatatcac	agacattcaa	177900
catctggcaa	atgtcttgac	aaaggagatc	acttgtgttt	gctgcagtgt	cccttctggc	177960
tgtgagattt	tgctcctcac	cactgcagga	ctgcagatct	attctgcctt	tttagttgac	178020
ttttcattcc	tgagaactgg	ggaaaactga	ctttgtatft	gggctttgaa	tttgtccatt	178080
tgtcaatcca	tcacaccaga	cctaaccaac	tgccaagagt	tctgctgact	tttgttttct	178140
ctagggtgggt	cacttttgctg	ggctcatcct	catccttggc	ctgcagtfta	tcccagggaa	178200
agaaaatggc	taacgactgc	taagaagcag	tctttccttc	cagaaatftt	agtctatcta	178260
gaccttgctg	cagtctgaag	tctttaaaaat	gtgtttgtta	tggtagaata	ttttgagttg	178320
ccttaggagt	attgcttgct	gtcacctatc	atattctatc	aggaagcaga	cgtcccattt	178380
accaaattgtg	aagaaatatg	gcatcaatac	ccactgcaaa	aagtgtaaat	aaataataaa	178440
aaaatagatt	tattacagag	tgcaagggaa	aagaaaaaaa	tcagccagtt	tcagaattgt	178500
aaactggacaa	atgttggtac	agttcatgaa	gaggttctac	aaaatggctg	gggggtgggaa	178560
cataatgagt	tagtttgctt	ttttttttct	ctttccttcc	ctttcctttc	cttacaaggt	178620
ctcatgtagt	ctatggctct	aaactcacca	ctgtaaatca	ccttgaactt	ctgatccttc	178680
tgacagctgg	caatgtaagc	atgtgccacc	aggcctggct	cacacatttg	gttttttcaat	178740
acagaatagc	tctgtgatga	ttaacttcaa	tcactcaact	gacataacca	agaatogtct	178800
gaggaagagt	ctcagtgact	gggtgggcta	acggcatgct	cataagggat	tatcctgatt	178860
gttaattgac	atggaaagat	caagtccatt	gtgagcagca	acacgccctg	aacagaagtc	178920

ttctgaagta	taagaggaga	aagcttgatg	agagcaagca	ggcaagcaag	ccaggatcca	178980
cggtgttatt	ctctgtctgt	tcttgaccgt	agatgtgatg	gctgtcttgg	cttcctggga	179040
aacatgaact	gcaccctgga	attgcaaggc	aaacaaacct	tttcctcttc	caagttgctt	179100
tatgctaaga	tattttatcg	cagcaataga	aatgaaactt	agaacaggcc	cataactgcc	179160
agctttggaa	ctgaacctaa	ggctgttata	attcactagg	atagggacca	ctggaagtga	179220
atctgatttt	gatggtaaaa	tcatgtgttt	gtttctggat	atgatagatt	tatcaatttg	179280
agactcagaa	aagaagttag	gacttgaatt	ccgttttaga	gacattccag	agaaaaactga	179340
tgtcattgtt	ctgaatgtaa	gtgcctcagc	tgaaaataca	aagagtacag	ggaagaaagc	179400
ccaggctaga	atctgaagga	actcctctat	tttttgtttg	cttggttgtt	tggttggttt	179460
tttgagacag	ggtttctctg	tgtagccctg	actgtcctgg	aactcacttt	gtagaccagg	179520
ctggcctcga	actaagaaat	ctgcctgcct	ctgcttccca	agtgtctggg	ttaaaggcgt	179580
gtgccaccac	accaggctag	gaactcgtct	attacacatt	aacacccttc	tttaattaac	179640
tgttctctgc	aatgtaccaa	atagtcaatt	gattcctggt	tattttaccac	atgtttctgt	179700
tagtaaacca	gaataactta	tctagccaaa	gtctgcctat	tagccatatt	ttcatcagtt	179760
cccaaccatt	tttggaattc	tgtgagggga	atccacagat	gctgtagacc	gcttttagaca	179820
tttttcagct	tttttcaagt	tgcaaggtcat	gattcagtggt	gtcatgaaat	taattttagt	179880
ggttcttgatt	agcattttcaa	aatgaggcaa	gcagagggca	tattgtcaca	gcacagcaca	179940
tgccgtaagc	agccacacac	tcttgcttgg	aggcttagtc	agtttctggc	tctaaacgcc	180000
ccaggtttgt	ttctctatcc	taggcctctc	tcttaaatcc	caaacatagt	tagacattac	180060
cattggggca	cgtgcaactc	aaacacggag	tgtgactcct	ttccccatct	gcggttccca	180120
gatttggtcaa	tgtcacccctc	ctcccttctc	cctaggggtca	gttttacctc	tcacactcca	180180
caacacaaca	cctctcatct	caagaattgc	cattagggct	ggtgagatgg	ctcagaggtt	180240
aagagccacc	actgctcttc	tgaagggttc	gagttcaaat	cccagcaacc	acatggtggc	180300
tcacaaccat	ctgtaatggg	atctgattac	ctcttctggt	gtgtctgaag	acagctacag	180360
tgtactcaca	tatatataat	aaataaatct	aaaaaaaaaa	aaaaaaagaa	ttgccattaa	180420
atgtacctca	gagtcctaat	gcttcttctc	cccctgacta	cactcacgct	ggcctgagtc	180480
cattttctta	ttgaggttac	tgcttctctg	cttctaccct	ggctccttct	gctgcctatc	180540
cttgacacag	cagacaagca	gttctttaaa	gcaggggtca	ggaccagtga	gactgatcgg	180600
ctctggtggc	acttctctgc	atgactgatg	atctaagggt	aagcctagaa	cccacgaggt	180660
agaagcaaag	gacctactct	ccaaagccgt	cctctgacca	ccatgtgtaa	actgcacatg	180720
tacatgcatg	cacatggtac	acacacatac	acagaagtaa	aaagagattt	aaattgaaaa	180780
tcattaaaaa	gaaaaatcag	ggctcagcaa	actttccgtg	tagaaaaacta	gagtaactag	180840
gctttgaaag	ccaagaagtg	gatattaatt	atagttattc	attatagcag	agatttctaa	180900
aaccttttga	caaaaactaaa	aaatataaca	gagtgatttt	tttttgtaat	gtaagtttac	180960
taatggcagc	agtgggatta	gtttcttttt	tagattattg	ttattatttt	tattaattat	181020
tagtggtttt	gtgtttattc	atattccaca	gcatgtgtgt	ggaattggat	ttctgcttcc	181080
acctttgtgt	gggtcctaga	gattgaactc	aagtcactca	gcttgcacag	taggttggtca	181140
ggcttacaca	gtagggtggtc	aggcttgtat	ctttggaagg	caagcatttt	acttctgtgt	181200
ccagctcact	ggccttcttt	gtttaaaaaa	agaaaaaaaa	agtccttttt	tggttaatta	181260
gggttcatgg	ccagtgtctc	ttatcttaaa	atcaactgca	aactttttatc	tggtaaaaag	181320
ccatccttag	ctgtgtgtcct	aggagaaaaa	catacagttg	gatggcttta	tcctgcaggc	181380
ttagtttgat	catctctctt	tgaagatata	atcagctcac	atcacactca	agcctctgcc	181440
aacgagtttt	ctacttctgt	tcaacaactc	accacaagct	agcagctcca	aacaacagcc	181500
agttatgac	ctcacagtcc	gggtgggtcag	aagcctaagc	gggcgtggct	acctcgctgc	181560
tattgcctga	ccctgctcgg	tgcatccaca	ttcacatcct	ttcctggtga	gtgtggttct	181620
ttgactgggt	ttgttccaat	tttttagtata	tgtgtctctg	aaacaatctt	tttgctctctg	181680
cctccagact	gcagggatta	atgttcttga	ctgccacaga	gcactaatat	ttactgaaca	181740
tgtgatcatg	tggtgtctcag	cactcttgca	ccgaaggctc	ggggaacatg	gaggaagagg	181800
gggtggaaag	attccaagaa	ccagagggaag	agaaaggtca	gaggtgagac	tgcatctcct	181860
agaaatgtca	gggacatttc	tagacctctg	aagtctcaag	aacaaggcct	gaaagtctta	181920
tttatatagg	ttaacctgaa	aggggaaaaa	attcttacag	gggtccaacg	ttagacaaaag	181980
aactctaagc	aactaaggaa	tggtgggggg	ggggtagtct	tccccaggga	acactcctct	182040
acccttcaag	ccccacccaa	gctggttatc	caaaacaaac	tggtcagtc	tgaagccata	182100
tacgcacaag	taacatcata	tggtggggca	gattgcattt	aggaatacac	acatacacac	182160
acacaactta	aaaagagagg	ccatgaattt	aagagagagc	aaagcaaagt	gggaaggggt	182220
acatgggaag	gttgagggca	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182340
nagagagaga	gcctattatg	tcgttggttg	cttctaatac	ctgtcccaga	gcaaggccat	182400
gctgagtcag	aactagccta	gctgagacac	tgctccacca	ctgtcccaga	gcaaggccat	182460
cgctgtccca	gatttgcctt	tgggggccct	tgaaatgaaa	gtcaccagca	ggctccggaa	182520
gctgcctcat	gctaatacgt	tgaggttggt	aaaatagccc	tgaggtgggt	cctggggcctg	182580
cagctggggc	agagccacta	aggggagctc	ggtccttttg	agcagagtta	acagtcatca	182640
gtgctttttt	ttttttttta	aatgttccct	gctttaggct	cagtgtctgtg	cgctacttct	182700
aatccttgca	ataggctgca	agacaggcaa	gaatatcatc	cctgttttgc	cctcaggcaa	182760

attctgaagt	ctggcaaatg	aaagatgtgg	gatttgaaca	cagacttggt	tgcccaaaga	182820
attctcactc	tacttctgcc	tgtgccacct	tcctctcatg	cacggggagg	ggaggggagc	182880
ccacctccca	tgctcagggg	ctaggaagtg	gggagaagat	ggatgtcctc	aaagcagggg	182940
gagaatgaag	tagaagccag	cttcaaattc	aaactaagca	atgtttttatt	tccattttccc	183000
tgaacataaaa	gttcagttac	atttgggttta	aaaaaaaaaa	tccttacaca	actgggtctt	183060
gagaaatgtc	aagtgtctaca	attcagtggg	tgtggatgaa	acaatcaaaa	tggtgaaacac	183120
ccccaaacag	atacaaacct	tcatcaaaag	ctcttccaaa	ggctgggtct	gaaaagagcg	183180
actcatgttc	cagcccagtt	ggctccttct	catgtgagct	ccgacttcca	aagactgctt	183240
gcaccaggag	gaaatataat	agatgtcctt	tttaaggggg	gtgggggtctg	tctgacaacc	183300
tcccacagtg	actgtggata	cagcccagtt	agtagagttc	tcgcctagca	agcgtgcggc	183360
cctgggcttg	agatctagca	ccctaaggca	tagtggatga	tgcccatgac	cacagcactt	183420
gggacgtaga	ggcagaagg	tcagttcaag	gtcagatgag	gggtggggag	gcattccttt	183480
aatgccagca	tttgagaggc	agaaacagat	gaattttgtga	gttcaaggcc	agcctgggtc	183540
acagactgag	ttccaagaca	gccaaggcta	cacagagaaa	ccctgtcttg	tcaggaaaaa	183600
agatagtggg	agagaattca	aggttatctt	ggactgcata	agactttgat	tccaaaataa	183660
acaaaaatgg	agcatgaatg	cttgcaactg	tggaacaatat	tggggttcata	catatttctgt	183720
tttgtcacct	acataccaat	tatacaaatc	agatttcagct	gggcccactg	gtgcatgttt	183780
gttcccagca	tctggggagg	agagatgggc	agagctctat	gactttaagg	ctagcctggg	183840
ctacaaagta	agttctagga	cagccaaccc	tacagagaga	tacactactt	ctaaatcaat	183900
caatcaatcc	atcagtcaat	catgggctgg	agagatagct	cagtgatcaa	tagtgccatt	183960
cttccagagg	acctgggttt	gattcccagc	acccacatgg	cagctcacag	atgtctgtaa	184020
ctccaatccc	aagggacatg	acaacttcta	ctgtgtctct	tggtcaacag	gcattgcacg	184080
agcatacaat	atataatgg	gtaaaaatgct	atatataata	aaatcagatt	cacaaatcaa	184140
gtacagaaag	agatcaacta	taatgaaaca	accataacac	atactgtttt	aaaagctacc	184200
tttctgtct	gacatctgac	ttcttgcgtg	acccagtgcc	tccacaatga	gacaagcaga	184260
tggtgtagtc	cctgtgaccc	aggattaggc	aaccactgcc	cctgggaagc	cattcctagt	184320
aacaggatca	aatcccacag	tgctccactg	taccccgcc	gaggacgtga	agcctccctt	184380
ccccggctgt	ctgtgcagca	ccgcccagat	gcttgggtgc	actgagtgac	catcggagcc	184440
caactgagga	gtgcctcagt	gctcctcagg	gcagtgtgca	attgaaactt	gcatgtgatt	184500
tctggaattt	gtcatttgat	atttccagac	tccagctgac	cttggaatac	agcacagagg	184560
gcagaactgc	agaggaaaag	gggcactact	gtgactgtta	tcccctgcat	gattatagaa	184620
gggtctgtgc	tttctcttac	aaatcattgc	ggcctctctt	cacttccctc	ctttgaagtc	184680
gaaaaaaata	gatgcatect	cacagtacag	gagctcaggt	acgaggcggg	ttccgggact	184740
gaggcaggac	tcatcatgca	gcctcttttc	ctcaactaca	ccgcccggcc	tgacagagtc	184800
cctctgatca	aatcagtttc	aggcctggaa	agaaacggcc	actcaggetg	gggatgtggg	184860
tccattagag	gcctgcatgc	acgaagctct	ggatttgatc	ttcagcacgg	gcataagccc	184920
agtatggtag	tatctgttta	gcattggtg	gaggtatacc	tatctctgca	caggagagcc	184980
agaagttcag	agttatcctt	gcacttacag	taagttcaaa	gctagcttgg	gcaacatgaa	185040
gttttgcctt	aacaaaacga	aacaagaggg	gccggggaga	tgcttcgctg	tgctgagtca	185100
tttgctgcca	agtttgatga	cctgagtttg	gtcccttgag	cttatggtgg	aaggagagaa	185160
atgacgcttg	aactccacgc	acatgccaca	gcccacgcac	gaatgtgtat	acacacacac	185220
actaagtggg	taaatgttaa	aaataaataa	atcaaaggaa	tgccactca	aaatctacca	185280
tcgttgggaa	gggaggggaa	aaggcaggcg	agggagatag	ataaccctga	tatgaacacg	185340
gaaagagcca	gtgtgccacc	aaagctgccc	agtgtgccac	caaagctgcc	cagtgtgcca	185400
ccaaagctgc	ccagtgtgcc	accaaagctg	cccagacttg	attacagatt	tgccagggga	185460
cacaggaggc	cagcaggagc	agccaggttc	cacctcagag	gtggagccac	aaacctggaa	185520
atgaaacgtc	tttccctttc	ttcagaccac	agcagtgaac	gctgtcctgc	agagtctgga	185580
gggctggcag	ggctcatcca	ctctagtgtg	cctgtggcca	gaacaggcct	cagtcacagg	185640
tgctttttcca	aggtcttagt	gtctaattaa	ggtttagcagc	caaattggag	agagaagggg	185700
gctggacttt	actctgctgt	aaggactttg	ggcattgttc	cattccgtga	tcaaatacca	185760
ctggctctgc	caaccacccat	gtcagtgggt	cttcagaggg	agaagaactc	atcctttttt	185820
gagaggtttg	gtctggtcct	tgtctaattg	aaaatgcctg	gggcaccagg	ttaatgtcaa	185880
ctcaaaggca	agtgtgtggt	cagcatgtgt	ggaatcctaa	gttcaatacc	catcagagcc	185940
ccaagcccta	gaggacagac	atgctttaaa	aaaaagtcac	gctttaaaaa	aattctgttg	186000
agggggctgg	agaaatagct	cagcagttaa	gagcactagc	tgctctttca	gaggacccag	186060
attccgtttc	taacattttc	atggtggctc	aaagctgtct	ataattcaag	tcctagaggg	186120
gaatctgttg	ccctctctgg	ttttctcagg	caccaagaac	acatgtgggt	caaacatata	186180
cgcaggcaaa	acactcatac	atatgaaaca	ttttttataa	acctgtctcg	atgtgggtggc	186240
tcatgccagc	gatgctctca	gcactcagat	ggcagaggca	gggggatttt	gtgttgagcc	186300
cagcctgagc	tatagaatga	gatgctgtct	caaaaagaaa	aaaaaaacaa	aaaaaaacaa	186360
aaaacaaaaa	caaactctgtg	ggcttaatac	ttcctagcca	gaaggagctg	gctccagcaa	186420
caatggatcc	ctagagctct	gctttgcccc	ctggtctgga	gagcttgctc	tagaaaaggaa	186480
ttctccatag	accacatttc	tattttgggg	accactctgg	gcctgcacat	ggaaaaatgg	186540
agagttgagg	tttcatgggt	catttttttt	gttggttacag	taatgagagc	tttgaaatga	186600

tcacaaaagg	aaaataggaa	aatatgcctc	ctaaaaagag	ccgaggcaaa	ttattatagc	186660
aaccgaatcc	taaaaggaat	gttcaagtaa	aaaaaaaaaa	atatggctca	tgcgaagttg	186720
ttatggcaac	tgacatttaa	aagtaacagg	atttgggcaa	gtttctttct	ccaccctctc	186780
tgcaaagtct	tgagggaact	tcatgctaaa	attatgcttt	aattttaatt	agccaaatag	186840
gtcataaata	tagcttattc	ccaaatcctt	agaatttcta	cctgcaagg	agctgaacta	186900
ctaattgagga	aatatattcca	caaaaaccca	cttaccatta	agaaccccca	tccgatattt	186960
ttctaataata	atttatcaat	ttaaacacat	tgcataatgt	gccactctgt	agcattccat	187020
taaaatgata	aatagcaata	gtggtgaggg	gtggggggca	aaaaccagga	gattaaacat	187080
atccaaagca	gtgtagctat	atttaaatac	ctcaatccat	tgtagaggga	aacacactgt	187140
tctcatccgc	agatacagtc	tagactcaga	gcagcatatc	cttgactgta	agggatttaa	187200
taggacagac	gaaggggggc	aataagaaat	gacaggaaac	ttcagaagaa	ataaaatttc	187260
tattaggctt	tgttataaga	ttacatcaaa	gcagttcata	tgttttaatc	tggggaggaa	187320
aaaaagcaac	tacttggggg	ttgcgcctgg	gggctgcctc	tgtgtactga	accagacagt	187380
ttgcataatg	aacaattttc	attcaatcag	gatctcagca	gagatagctc	ctactcaaag	187440
gaaccgggca	caggctcata	gtttttatct	cccagctcca	cctgctggag	aaaccttgta	187500
ttgcagggag	agaaagcagt	cgggaggtct	gtgtcctagt	gctgtgtacc	taaagtata	187560
gacctgactt	taaacagttt	ctctctggag	gttgaaaggg	gctctgtaag	ataccagagt	187620
ggattgctct	caaagactct	cggactcctg	ttacaggcaa	gtaaggctct	agcagatggg	187680
agcatggatc	tccggccctt	ctcactgctt	tcttgaatca	gggatttaga	aattgctatt	187740
tgcataccag	gaggactgaa	gtttggctcc	cggtgaccag	aggacaaggt	cattgtttta	187800
aaccacccaa	actcattttc	gacttgggtg	gtcaaatctt	caagtttccc	agcagtctaa	187860
ggattcataa	aataaggcag	aggcagagaa	acggagggtg	tgtgtgtgtg	tgtgtgtgtg	187920
tacccaaaat	ggaactgcat	tttcatgcac	aatagaaaac	ttaaagactg	aaccaatcat	187980
tttgaaaaac	tggcacagct	gacattggct	agagggaagg	acggccaggg	cgagccagct	188040
gcaccaagac	ccagggtgta	ggcctaatcc	gcctttatcc	gagggtttag	tgaggctccc	188100
gcgcctcacc	aatcccggct	ggagccgcag	aagagctctc	ttcacttggc	tcagctcccag	188160
cacagtcgca	ctatgctctc	ccgtggggag	gcccgtccgg	gaggggggag	gacatcaagc	188220
tttgtgaaac	tgttttcgaa	aacctgggat	gatcatttaa	atgtttaaaa	tatgcacatg	188280
gtaattcaaa	actaattacc	ctgagcacat	ttgaaacatt	tatgccatca	tcttggatcc	188340
tgcctactga	ttgtgcgctg	cagctcactc	tggtgtttct	ataaaactgt	tcagcgattt	188400
taacttccag	gctaaatcag	gcagccacag	gcgctgcctc	cagccctggg	ttgggtggaga	188460
gacccccatc	cctgacttcc	aggcgaggag	ggcgccgctt	tctccagaga	ccggtttgtc	188520
agggtcttgt	agttctggct	gccgaattat	tgctcttata	cgtgttcata	attctcatct	188580
gcattattta	atttaggcta	gaatgacctc	tttccctccc	gagtcttccc	ccctcattcc	188640
catttccctc	tcttcaattc	gtggccccca	ttttctgatt	ggtccaaata	tatagacaaa	188700
tatecttgat	cgtcccaccc	cacttggtcta	catcttcate	tgggagccaa	tgtgggtgagt	188760
tttctgggtt	tgcaagggtg	tcagggtccac	cagtcatcct	aagggtgtgtg	agagaggtag	188820
accaacatga	gcgggcgaca	gccgccatca	ctgagagagc	acgtgccctg	cagctcaggc	188880
acaggcatgc	acacaccggc	agacatgtgc	acatgcgctt	tccccagcaa	accctgcttg	188940
cagagtaatt	aggcctaggc	agttcctgaa	gcaaattcat	ttcccccttt	tccagaataa	189000
aatgagttct	cttccttttg	gggtgctaaa	ccagcatgcc	agtggctaga	agcctgagat	189060
gggtgatgtg	gctgaaacca	tttctgcagc	caagcctgtg	ggcagaagct	aaccttgggc	189120
tggggagctg	cagtcggaag	aggcacaatt	ctgggatcaa	gaaatgagca	ctgggtttata	189180
ggtacactcc	cagaaataga	cagatgaggg	ctgcctcctt	attaigcgtt	tgaagatgcc	189240
catggcgggt	ttttagacat	ttaggaatat	aaaagtaggt	tggattccca	cagtcagctg	189300
aagtttgaca	gagtgatatt	accgggttta	actagagcca	ttaagagact	cttcattatc	189360
ccacaccacc	gccacccaag	ttatcacatg	agccataatg	caagagaatt	ttcattccat	189420
caacaagaga	gggagccggt	ctatctttgt	ccaaaggaaa	tgagcagccc	agcgtgaagc	189480
ttgtgaggaa	ttgagtgtac	aacactccaa	taacatcccc	tgcaggattg	cctctgcgat	189540
ttagtccggt	aagcaggggt	aactgcgctc	gagcagctct	cctgtgtacc	tggcttgcaa	189600
gaacaccagc	tcgagggaaca	ccaaaaaggg	cgattaatga	caaaggacac	tcatagaggc	189660
ccgaattcca	cagggtctaa	gtattaaagg	ccaaagaaat	caaggtctag	gccatttctc	189720
tggcgctcag	caatctcatt	tattatttct	ctacaaagat	ccaacactca	atttcccagg	189780
tatccccctg	atctgactca	cattctcctg	ctcagtaagc	catcctgggt	tgaacggggc	189840
ctccccctcc	cctgcctatg	catgctttgc	gtcttcacaa	cgacagctgg	taatttgcaa	189900
gacccccctc	actggactct	ctcaccacc	atacttgtaa	ctactccttg	gaactacttg	189960
tttatcaagt	gttctgttgg	tgagccttct	cttgcattaa	agctgtgaga	aggaaccaca	190020
gtttctattt	cctttacatt	tctttagcgt	tctcacatgg	gagacaccca	ggttagatat	190080
actgaggggt	ctggtagttt	tagagttgga	gttagatgac	ccagcaacat	gccttcccc	190140
accacgcacc	aagcaaaaat	tgcacccacc	cttccctcag	atgttcctgg	catcttataa	190200
ctcgcccaaa	gccagattta	ttgctcctgc	tgtaaagtgt	atcttctcta	agcctcactt	190260
aaaagctacc	acttggcaga	agatcaagtc	agaagtgcag	gctagcaggt	gacggtgagg	190320
acagggcggg	atggggcggg	taggggtggg	cgaaataattg	aagctccaag	agttaccagc	190380
tcaatatatta	acctaactgg	taattttgctg	tgacaattac	gccatgaagg	gaacgctgcg	190440

actatgcaag	aatgttgctc	tctaattaag	agggctctgc	atttcctagt	caccgcgact	190500
ttaataaacac	acagaatgag	ccttggtctcc	gggagctaaa	ggttccatta	ggagcacggg	190560
cagcatatgg	ctgtgcacat	aggccgtgag	tgatgcagcc	cagttaagcc	cgctaacacc	190620
ttcaattcgt	cctcagatag	agcccagaga	gcgcggtcca	ggccctcacg	ccacgagccc	190680
catttgactg	acaggcatct	tcccggaaag	cctgcgcgtg	cctacactgc	aaatggacct	190740
gcttcccaca	gcccggcttt	caaccaggaa	ggcttggcgt	gggtctgatc	cttcaagagt	190800
aactttaata	aggattttct	cacagaaaaga	aaagtccatg	ggaacaaatc	ctcctcttaa	190860
gagcgtgaga	caggaatggg	gacacaagcc	aacaccccaa	ttgctaggct	aactctgata	190920
tgagacaaaa	gaatattaat	atcttggtcta	tgaaggagga	tggtgccatc	ttctgaattg	190980
atgggagttt	tgaggcatgg	ctaagctggg	caaaccattt	tctttttttt	ctcttcttaa	191040
ttagtggttc	atttatggag	ggcttgctgc	ccggagagcc	catcagaaga	gagctcgctt	191100
tatggagatg	tagcttataa	aactactcag	attttaaaaca	aacagtgcag	gaggccagag	191160
gtagaagtgg	tggggggtgg	gtggggcaag	agaacaattg	catctgcaga	aggctagccc	191220
tgacccccaa	gcctatgttt	agggttgatc	agcttcccga	ggcaagccca	gaagcctcta	191280
aaattttagg	ccaatagaaa	tgacctctgc	accacggctg	actgaagcta	taaataagcc	191340
tcgagtttag	cagtgggtgc	aacggagaga	gcagagggaa	gtccaatcag	agcttcatth	191400
ttttttttta	aagtcacatt	gcttgggact	cacctgaagg	cagggcattg	agtagagcct	191460
tggctccctg	cagcgagagg	ctccagtttt	cccaggcacc	agcccatcgg	ttggttacct	191520
aaccaccgaa	agggaaactgc	acagcacaca	agttaaatat	aggctgggtt	atctgcattt	191580
tacaagctct	gagcaagcta	tctgaagaag	ctgtcatttt	taatgacggc	acaaacttcc	191640
aattaccgac	tggttaatcc	actagggagc	aggtagtttt	ggaagaacag	ttcaccatta	191700
ttaaaagttt	acacaatcac	ttttgagttg	actataagta	tttcacacga	ggcaggtggg	191760
attagggact	ttttgggtgg	tttaactcag	gctgcaacca	acaatgagtg	ttttctcaag	191820
aattatacat	tgagattttg	caactgctgg	ggagtagtgg	agggtcctgg	taatgcagaa	191880
aggttatgaa	atggccagggt	aaggttgggt	gcttccaagt	ctcaaataata	ctcctaaggg	191940
cagctccaag	tcataagctc	aaacaagtct	tcaaggggcc	tggagagtta	agacaaataa	192000
ggatcactta	ggctacccac	ggacaagcac	ttctcataca	aggaccggct	acctccaaca	192060
ccatcttccc	aacatggctt	ctatgttgct	tcaacaacca	gggcaggggtg	aattaggggt	192120
gggtctctcc	aatgtggact	caaatacatga	ctacagcntg	gggttttttt	tttttttttt	192180
tttttttttt	tnntttggtt	ttcgagacag	ggtttctcca	tatagccctg	gctgtcctgg	192240
aactcacttt	gtagaccagg	ctggcctcgg	actcagaaac	ccgcctgctt	ctgcctctgc	192300
ctcctgagtg	ctggaattaa	aggtgtatgc	taccacgccc	ggccgagtc	gtcttgataa	192360
tgaagttccc	agtgcctgg	atgtcaactg	aagttggatt	ttactgtgat	gactactgag	192420
tccggctcag	aattttgggg	ggacaaggta	ccttgattta	actgggcact	acacgactgt	192480
aacccccaca	ttgggagagg	cagaggcaga	ggcagaggca	gaaagtgggt	tggaggctag	192540
gcaaggctac	acagcaagaa	gctgtctcaa	aaccaaagac	atctttcttg	atccaaatcc	192600
tgtcggaggg	tgtgaggcct	tggggggccag	aacaaggtgg	tcaaggaaga	ccactgactc	192660
tgtcctttgc	tccattactt	aatcagaatc	gccatcacag	atatagctag	gagattttta	192720
gccttggtgg	ctgcaatctg	catttaagag	ctaagtggga	taaactcagg	ggtgggcccc	192780
atgcctccct	ccccaccctc	cctgcatccc	tccatttacc	tgtttccagg	gatctgctta	192840
atttacctgc	cagccttttg	tgggacacag	gcttagtgge	ttagcgtg	tcggggcacc	192900
agagaccctc	acagaagcac	ctgaatgtac	tttcagcgct	gcagagcacg	cacggctcag	192960
gccatcaga	agaacccagg	cttatgtctaa	ggagccagaa	agtgaagca	ctgcaaaccat	193020
gtgattcagc	cccataaatt	tacacatccg	tacagccaaa	cccacttgaa	gtgatccaga	193080
gccactttta	ttgaaataga	aaagatgcct	attctggagt	gctaagtgg	acaggagggt	193140
gggtatataa	gagataatcc	catgttgtct	ttgatgtgg	gctagggaga	taaccagga	193200
cctcacgcct	gcctgcaagg	tagccaccaa	gccacaccca	caacctctat	ttatacacac	193260
actaagtgtg	gaggtatgga	taaaaaaa	tgtcccaaga	cctcacgaat	ctgcaaaccat	193320
ggtgcctgg	tggtggcacc	gtttggggag	gcagtggacc	atttgggtct	gcaggaggaa	193380
gttatgtcac	tggttatggg	ctttgagagt	ttgtagcttt	gctccccttc	cagttaactc	193440
tgtctctgta	aggttcctgc	caccatgttt	cctctgccat	tatggacacc	tggtcctcta	193500
gaactgtaag	ccacttactc	tcaggctctc	ttcagtcctg	gagtcttatc	atagcaatga	193560
aaagtaactt	gtgtggcagc	cagctaagca	aggctgtgg	ccgactgctt	gggattatgg	193620
ttgtgtctgt	ctgtctgtct	gtcattccat	ttatatgtc	ctgagaattg	aaccacttta	193680
ccactgacat	gtctcagtc	tcttggtatc	atataattcac	ttaagacaag	atctcattaa	193740
gtcattcaga	ctggttttga	gcttgcaatc	ctcctgcctc	tgctcaagg	cgataggatc	193800
cctagggtag	tcgaccagac	tgggagtagc	aggttctgtt	ctcttagctt	tctacagtga	193860
ttgtggatta	tttgtgtata	aagatctgat	ggcccagccg	actccccttc	ctttaagtga	193920
acatcaacag	tatttagcat	caacttaata	aactcatttg	gtaaagccat	ctccccacct	193980
cttgaacaaa	tgaaaatcaa	acagcagtac	ctgttctcct	agagcagcgg	ctctcagcct	194040
tccggccttt	taatacagtt	cctcgtgttg	cggtgacccc	cccccccccc	agccgtagaa	194100
ttatttcatt	gcttaaccag	agttaactgg	aaggggtaat	aataaaaacca	gtctgggaga	194160
ctaaggttac	ccaaccacgc	taggaaggag	aggaaagggc	cactcgca	aacctgtctt	194220
tgagatgaag	aacaatcaac	ataacaggga	cagagcagtc	cttgtaacaa	gtgcaaggga	194280

gagagagagg	ctgagtttct	acttctataa	ataaaccctt	ggcaggcgga	tcactaaagg	194340
aacacaagtc	aatataaacc	tttagacatg	gggctgccaa	acttcacttt	tcgacagtat	194400
attaattatg	tagtcaatag	ccatgggttt	cattagcgta	ttaaatacca	cgatcaatat	194460
tatttatact	tttcgaagac	aagccactca	gggaaaaaat	ggtgggggga	ggaggaggaa	194520
caatttgacc	ctgtagttca	aaaaaagtca	gaacagcaca	ctagagatta	gcaagggttt	194580
aatggaaggc	ataaaacact	ggaaatatgg	acagaaatca	gatccctgcc	ttcatttttc	194640
tgccttttac	aaagagactg	gagggaattc	agaaactatt	taaaataaag	gcaaaatgat	194700
tagagcccc	ccctcccctc	agctgcttaa	cactgggggt	gtgggtggacg	caaaataagc	194760
attgagctct	aagtgataga	tgagaatcay	aacaggaaca	gtgtttttga	ggcaaaatat	194820
gtccaagaga	attcaaagaa	ctgtggggcca	gaatctactt	aggcagtcct	ctgggaccgc	194880
aatccctcac	aggcgttaac	agtggaaacca	atttccaagg	cagccctgct	ggtgatctga	194940
tttttgagta	gggaaatctg	ttaaactctc	tcccacgagg	gagcccgact	ctttcactcc	195000
ccacgggttt	ctacatgcag	ctgtgctaga	tctgtctgaag	tggccggtga	ggaggtgtgg	195060
ggattgggtc	agcgacctca	gaggacattc	ttgttcacta	gccctcgtgc	actggggcga	195120
tgaccgaatg	ctgtgagcag	gagatatcaa	aggccggcta	ctggactgaa	aactagatca	195180
ccatctctaa	cctgcaattt	gtcaatctca	gacagcaatg	aagactgtga	ttttctagtc	195240
aacgctttgt	aagcaaggtc	agatagaggc	tccataaaaa	ttgttcaggg	ttcaggcgaga	195300
gaatcaagtg	taactcaatc	cctatctcct	gagattaggg	aagggaagga	aggctgtgtc	195360
tactaaacca	gtgagcctca	agcaaagcct	gtctgttctc	agcaagggtga	gccacccacc	195420
aaagatgcc	acagctaagg	gccaggggat	tagtgcaggg	tgctgtgata	tcaacagctg	195480
ggagacagaa	acaggaggat	caggacttca	aggtagtttg	ggctataaaa	tataagcttg	195540
aagctaccca	cttgaagact	gtccccaa	aaacaataaa	gctgggtatg	gtggtcgatg	195600
cttgtatttc	tagtgtatga	gacgaaggaa	gaaagctcaa	gcccagagacc	tgcttgggtt	195660
acatagggaa	gatgggtgct	caaaaacagg	acagccgagg	agcagacaga	cagggcagac	195720
agggtgcctc	gatctaaatc	cacatacctg	gatttaaagt	aatatctggg	agactgggtc	195780
gtgagggccg	ttccagagat	ttaataaaga	cccacctga	ctgagtatgg	gcaacaccca	195840
tgggtggccc	aggggtccag	actgaataaa	gggaaaatgg	gagggaagtc	agctggtagt	195900
gtttccaagt	attaggacca	cagcctggcc	cctggcatgt	gctggccagc	tagtctagct	195960
ccagtatcaa	gcttcaggcc	agcggcaggg	cactggacag	ttcccacaca	cgacacacac	196020
acacacagag	cactagcatt	cacctcctgg	tctcttcttg	acaacagata	aaatgtaact	196080
ggctgccaca	gtgagagtcc	cctaccttcc	tcaccgttaa	ggatggtaca	aactgtgagc	196140
cagcagcagc	catcttctca	gtaacttgct	ttccacagat	actgtttatg	cactaagaaa	196200
agcaactgaa	acatgggtg	ctgtgacccc	ttggcaccac	aaagccatgg	caagctgaag	196260
tgacacatc	acaggccagg	cctgaagatg	ctgggggact	gcaatgctgc	ctggattctg	196320
gcagagatgt	gcagcagatg	ccaagagggtg	ggctgcagca	accagagata	attaatatga	196380
ttaggaacac	actgagcagg	catgctcttg	ccgaatgaaa	agcctcgag	tgtaatgact	196440
gttttcttcc	tcgatcacgg	tctccacgtt	tcagagttgg	cttggtgtta	ggctgccgag	196500
taaacatcaa	tccaaccccg	aggggccaga	tcctcggtgt	tcctgggctc	aatcgctttt	196560
ccttttgtgt	tttcattcat	ttaaagatgc	attccagggt	tgcaaacatt	agtgagaatc	196620
atctccaggc	ctcagtcata	tctctgagtc	tgtaatgagt	taacatcttt	ccctagtga	196680
tatttattat	gaaggctaat	taattgcttt	ccagttacaa	gaatccttta	cagtcaaaga	196740
aagtaggatc	cacaaagata	tactgtttat	tcaaaacaa	caaaggaaac	aaagtcttct	196800
tcttaaatc	tatttaacat	agcttttaata	aaggtaacaca	ggctcgctcg	gcaaccgaac	196860
ggtaactgat	gcaaaactgaa	gccatgctct	gtagcagcct	ggatgtccca	gtgccacctc	196920
tgtctgcagg	ctttgtcgga	tttactaaga	ttctgttatc	ttcaaacagg	gatttgtgtc	196980
caagtaactg	accccactat	gtggataatg	aagtaaatta	tgcaatttgg	gggtttgctt	197040
ttccccagg	ggacagcaag	ccagtgccta	tcagccgtcc	tcagaggaga	caattctgat	197100
taatatcaga	gtcatctgac	tcagtctatt	aaacctatca	aacctgaag	gaaggatatt	197160
cagatatata	cgataggcct	ttgattaata	attctacctt	gttgccattc	taagcattaa	197220
caaccatgca	gtaactctgc	aaaacagacc	ctttgattcc	aggcagagcg	accctctgaa	197280
cacctgggtt	ctccccctact	cttctcccc	caggaggaac	tcaagacaaa	aagggtgccac	197340
cactggaaaa	gcacactcca	ggttacataa	tttgcctcat	tatccagagt	gggtttaatg	197400
actgtgaca	taatttctgt	ttgaagataa	caaaaattca	tgaaatccga	caaagccgga	197460
aggcaggagg	aggggactgc	tgccacacta	ccggtggctg	agaactggag	cggaagggtc	197520
acacacagcc	ctctgagctc	actgtctttg	cttatcagtg	agtcccaaga	ggggcccaga	197580
tgggttgcca	gcctccccta	gaggatcttc	attgtggagc	tgtcccatgg	ggcggaagg	197640
aagccattct	atttctgttc	ttctctcttc	cgttctggcc	accagtggta	cttgcctcca	197700
tcacatgttc	ttcctgatgt	tcgcgatcag	cgtctgcca	tagtctctga	agtcacaggg	197760
cttcacgtcc	atcacagtgg	ccttaattcg	agattcatcc	tttagaaaag	agagaagctg	197820
tttgtgagt	gcagagcctg	gcgtgcagcg	gaagagagaa	ctttctttgc	ttcagtggtc	197880
tcaatgagtc	cagcaggaag	aaggaaagt	tacaagtctc	agagagaaa	tgctgtgact	197940
tectggagtt	gggccagatc	ctcttcacac	gaccttttcc	ccatcctagg	tgccctgtgc	198000
tcagacctag	catcctccc	gagaagcctc	tgcttttcta	tgggtgcagt	gggggcccag	198060
agcagacagg	taactcacc	taaagcatca	ctttcatcta	gaggagctct	gtggtagtag	198120

ggactgaggc	ttctgctcca	gctctgggca	aggttacttc	tctgctcttc	accattcctg	198180
tccatcccag	gaagacagaa	aatccctaca	ctctcccttg	atctaccoga	ctttctgaca	198240
ccagcctacc	tatgttcatt	taatacaaca	actaaaatat	ctattcacag	gcactaagct	198300
ggtgataacg	cagaatgcac	aaactctgcg	gctgcagggg	agacggcaga	gttccctctc	198360
cacttgtctc	cttgaactaa	acagtgctct	tgaggcagaa	caggggtgaca	cctaggggaca	198420
cacaagtcta	gctggggggc	ttcatgcttc	catgtgctta	gtaattaatt	actacatgca	198480
ccgctgttta	caagtatggt	taggagcccg	actgcctggg	ttggcctctc	gcctctgcca	198540
ctccatggct	ttaggttcag	agtcattctc	tgcctgcctc	tgcctgtctc	tccgttggtg	198600
aagcttgcaa	caacagctcc	aacacagaaa	gtgctgtgag	ggtcgacagt	ggatagatgg	198660
ctagatagat	ggggcaggac	ggactgtcca	gtaagcaggg	ttcatcatgg	ctatgcagct	198720
ctggacatca	ggattagttt	aaacacttgt	caggtggggc	actttttacca	gcacgtgcta	198780
tttgtttaat	attctgagtt	ttagaacctt	aactgtggga	aacaagagtc	cacacataac	198840
annnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198960
ctcctgctgg	ggttacaggc	ttgagccacc	atgacagctt	tagcaatagc	tttgtaaatac	199020
cacagtgtca	agctggatat	gatggcacat	gcttgcaata	ctaacctcca	aagattccct	199080
gaactcgagt	tggtgaaata	gtccaccagg	taaaagagct	tgctgccccaa	gcctgaatct	199140
gattccctgg	tcacatgctt	taaggaaaga	acttgccgaa	gatgtcctct	gagtgccacg	199200
tgtaccgatg	catgcttgca	gccaccacaca	caccacacaca	agtgcactgt	ctcacacagt	199260
gagaacagca	agtgaacaaa	caaacaagcc	gggggggggg	ggattgtgac	cagaataatt	199320
gagggggggg	gtaaagctct	tggcaggtgg	ctggctcctg	gtaaacactcc	ataagtgagg	199380
aagttccaca	tgtgaagtca	tgtgatcgag	tacatctggg	cctccaacag	ctcttgngga	199440
agaaacagat	gcagtctgtc	atatctaaac	cattgttgtc	gtatctctgg	gtagtctttc	199500
ttttctcctt	cctttctttt	ttctctccct	ttctctttta	aaaaattatt	tattttattat	199560
tatatctaag	tacactgtag	ctgtcttcaa	acacaccaga	agaggggtgtc	tgtcagatct	199620
cattatggat	ggttgtgagc	catcatgtgg	ttgtctgggt	ttgaactcag	gaccttcaga	199680
agaacagtca	gtgctcttaa	ccgctgagcc	atctctccat	cccccaacct	ctttctcttt	199740
tgagttaggt	tttgtgtagc	cctgggtggc	gttaccttaa	ctacactggc	tttgaacttg	199800
caatgatact	ctgcctgatc	tgtcttaatc	attttgagat	agggactcac	tacatagcct	199860
ttgctggctc	ggaactaaca	gagatctgcc	tgtttctgcc	ttgcaaatgc	tgggaataaa	199920
gttatgtacc	accacacctg	gagtttaagg	gttttttgtt	tgtttgtttt	tcgagacagg	199980
gtttctctgg	gtagtccctg	ctgtcctgga	actcgctctg	tagaccaggc	tgcccttgaa	200040
ctcagaaatc	cgctgcctc	tgccctccaa	gtgctgggat	taaaggcggt	tgccaccacg	200100
cccagtttaa	gggttttttt	ttgtttgttt	ttccttgaga	caggggtttct	ctgtgtagct	200160
ctggatgttc	tggaactcac	tctgtagagc	aggttagcct	tgaactcaga	aatctgactg	200220
cctctgcctc	ccaagtgtct	ggattaaagg	cgtgtgccat	cactgcccag	tgattttttt	200280
ttttttaaag	tgtgtatttg	tatgggtgtg	tggtgtcttg	tgggaagccag	gtgtcagatc	200340
cccagagcgg	aagtgttctt	aaccgctgaa	ccatctctct	tccctcttcc	ctaaactctga	200400
ttttaaaggc	accaaactct	taggtaggag	actatacaca	cacacacaca	cacacacaca	200460
cacacccgta	cacacccgta	cacaccacat	gaccatgcct	gagcacacaa	gtgggttttat	200520
tgtgtgtctg	gcctgtgtat	gagctggaac	caaaaccttt	gtcgggagat	ccgcagtctg	200580
cagtttgagc	acaggctctc	tggtttctgt	tctctgtcct	gtgtcgcatc	ttgactagag	200640
gcagagaagc	atctgcaagg	ctgtgaccac	gctgtgctgt	gctctgccat	ctacattttc	200700
aacaggaaat	ctcaggagag	tatttccttt	taagaacgcc	agacttttgt	gcctgggcca	200760
cttctctact	tcccagaaca	ttgtgtgcca	agtggcaagt	tattaaccaa	gtgcttttga	200820
aaattaaact	ccttggtttg	cagagtagca	tgggagcatt	gagaggggtg	atgcctaaag	200880
gcctggttct	gctgctggca	gagctgacac	ttggctaaag	ggctggcatt	tctgagatga	200940
gcctcactag	atccgcgtct	cagagtctgc	aggagaaatc	agagagggga	gaagggtccag	201000
tggcctgttc	aggatgatct	tcctctgcat	ttaagggcgg	ctggtttgcc	cacgtagccc	201060
cagaaccaa	cgagcctcgg	acgaagcccc	ctaaaggcag	taggagagac	tgagccttgg	201120
ctcttcagca	gggggtggga	caagagcaag	aggcgggatc	tcgcccggcc	cttttagagac	201180
acgtgcgggt	gtttccgtgt	ctgggagatc	acatgaccgg	catcagctga	cccgtcacgg	201240
tggagctcag	cgctggtgct	tcgcgctccc	cgccctcgct	cgccccggag	cgcaggaccc	201300
tgcggagggg	taagaaaacc	cccaggcttt	ctttcctttg	tcgctgggtc	gcgcagtcac	201360
ctgcacccta	cccccgctc	ctcgttcctc	ccagtcttcc	cggcctggca	ccccggaagc	201420
cactgcgagg	agggccgtgg	ccaggctcag	ccttgcgctg	ccccaggcgg	gccaggacca	201480
aatggcccag	gggagcagaa	ggcggaaggt	ggttctttaca	gcagggtccg	agggctggct	201540
cccttctcca	ggacctgaca	tggaggagct	gtgcccggag	gtggagagag	atctgaacat	201600
tgatgcccgg	cagctggccc	tggcgccggg	gggcactcat	gtagtggccc	tagtgtccac	201660
gcgttggtctg	gctagtctcc	gggagcgccg	actgggaccc	tgtccccggg	ctgagggcct	201720
gggtgaagca	gaagttagga	ctttactgca	acgttcggta	cagaggctgc	ccccaggctg	201780
gactcgagtg	gaggtgcatg	ggctgcgga	acggagactg	tcctaccgcg	tgggtggagg	201840
cgtgcccttt	gaggaggggt	cctgtagccc	tgaactctc	actcggttca	gcaggaggt	201900
ggctgcccag	aattaccgga	acctgtggcg	ccatgcatac	cacacttatg	gacagcctta	201960

cagccacagc	actgccccct	cagctctacc	tgccctagac	tctatacgac	aagctctcca	202020
gaggggtgat	ggatgcacct	tcttgccagt	gggtgaatcc	atcccatgtc	tatcaaagt	202080
cagggatggg	cctgccccct	ctcggggcag	ccctgcectgc	cccagccttt	tgcgagctga	202140
ggctttgtgt	gagtcgcccg	agatgtctcta	tgtggtacac	ccttatgtgc	aattctccct	202200
gcatgatgta	gttaccttca	gccctgccaa	gctgacacac	agccaagcca	aggtgctctt	202260
tcttctcttc	cgtgttctga	gggccatgga	tgctgtcac	cgccaggggc	tggcctgtgg	202320
ggctctgtct	ttgcaccaca	ttgctgtaga	cgagaagcta	tgcagtgagc	tccggctgga	202380
cctgagcgct	tacgagatgc	cttccgagga	tgaaaaccag	gagggctctg	aagagaaaaa	202440
tgggacaggc	attaagtctg	aaaaagaggg	ggaagggaga	actgagtgtc	ccacctgcca	202500
gaaagaactt	cggggccttg	tgctagactg	ggctcatggc	cgaatcagca	acttccacta	202560
cctcatgcag	ctgaatcggg	tggcaggtcg	acggcagggg	gatcccaact	atcacccagt	202620
gctgccctgg	gtgggtggact	ttaccacacc	ttatgggctg	ttccgagacc	ttcgtaaatc	202680
caagttccga	ctcaacaagg	gagataagca	attggacttc	acctatgaga	tgaccocggc	202740
ggcatttgtt	gcaggtgggtg	caggaagtgg	ggagccaccc	catgttctct	accacatctc	202800
tgactgtgct	tctgacatca	cgtactatgt	atacaaggcc	cgtcgcacac	cgcgtcgggt	202860
gctctgtgga	catgtccgag	cgcagtggga	acccacagag	tatcctgcca	cctctggagc	202920
gatgcagacc	tggacaccgg	atgagtgcac	acccgagttc	tacacggacc	cctctatctt	202980
ttgctctatc	caccctgaca	tgcccagact	ggatgtgccg	gcctgggtgca	gttctaacca	203040
ggaatttgtg	gctgcccata	gagccctcct	ggagagctgg	gaggtgtccc	aagacctgca	203100
tcaactggat	gatcttacct	ttggctacaa	actccagggc	aaagaagctg	tgaaggagaa	203160
gaatgtgtgt	ctgcacctgg	tggacgctca	caccatctct	accagctatg	cgcgtgtaca	203220
gctatttgat	cagccacacc	cccaacgcct	ggctggatct	cctgccctgg	cccctgaacc	203280
tccactcatc	ccccggctgt	tggtccagcc	tattcgggag	gccacaggcc	aggaggacat	203340
ttcaggacaa	cttataaatg	gtgcggggcag	gcttgcgtga	gaggccactc	catgtgagac	203400
tggctggact	agagataggc	ctgggacagg	agaagatgat	ttagaacagg	ctacagaagc	203460
tctggattcc	atctccctcc	cggggaagc	aggtgaccag	ccaggctctt	cctccagtca	203520
agcatcacct	ggcctgttgt	ctttttctgc	accctcgggg	tctcgaccag	gccgtaggag	203580
caaagctgcc	gggttggacc	ctggggaggg	tgaagagggc	aagattgtcc	ttccagaggg	203640
cttcagctcc	atacaggcct	tggaaagact	ggagaaagtg	ggtaaacttc	tggccaaagg	203700
cctagggagc	cagttggagg	agcctgaaaa	gcctcacgcc	cagccacctg	tgacacctga	203760
gagcctcttc	catcgagaca	tgacgtcctt	gggtgtcctg	ttggctgaga	tggtgtttgc	203820
caccagggtc	cggatactgc	agcctgatgc	acccttgtgg	gtacgctttg	aggctgttcg	203880
gggtctctgc	atacgccact	ccaaggacat	ccccgtgtct	ctgcagcctg	tgctagacac	203940
actcctacag	ctgagcggac	ccaaaagtcc	catggtgtcg	aagaagggca	agctagaccc	204000
actgttttag	tataggccgg	tttcccaggg	attaccccca	cccagcccag	cccagctcct	204060
cagcctcttc	agctccgtgg	tccccttccc	ttccatactt	ccagcactgc	ccaagctcat	204120
tcttttatat	caggcccggc	gtgtggagga	tgaggtccaag	ggtcgggagc	tggcgtttgc	204180
tctgtggcag	cagctgggtg	cgggtgttaa	tgacatcact	cccagagggt	tagagatcct	204240
cctgcctttc	gtgctgtcgc	tcatgtctga	ggagcacacg	gctgtgtaca	cagcctggta	204300
cctatttgaa	cccgttgcca	agggcctggg	ccccaaaaat	gccaacaagt	acctcctgaa	204360
gcctctctac	ggtgcctatg	agagcccctg	cgcctgcact	ggccgcttct	accctgtacac	204420
cgaactgttt	gtggcccagt	tggtgggtcg	gctgggcttg	caggcccttc	tcacccacct	204480
gctgccccat	gtcctccagg	tactggctgg	ggtggaggct	tcccaggagg	agggcaaagg	204540
cctggctcgg	accactgagg	atgaggaaag	tgagctcccg	gtgtccgggc	ctggctcctg	204600
tgcccttggg	gaagagattc	agatggatgg	gcagccggct	gcttccctcag	gactggggct	204660
cccagactac	aggtcggggc	tcagcttcca	tgaccaggcc	gacctgccgg	acacggagga	204720
cttccaagct	ggactctacg	tggtcgaatc	tccacagccc	caggaggctg	aggccgtgag	204780
cctggggcag	ctgagtgata	agagcagtac	cagcgaagcc	tcccaggggc	aggagagggg	204840
tggggatgat	ggcgggtgcc	ctgcggacaa	gaacagcgct	aagtacaggg	acagcagcca	204900
ggacttgaag	cagagcgaag	gctctgagga	agaggaggag	gaggaaggct	gtgtgggtgt	204960
ggaggaggag	caggaggatg	aagtcacggg	aacatccgag	ctcactctgt	ctgacacgat	205020
gctgtccatg	gagacggtgg	tggtcctctg	tgatgggaga	gacagagaag	aggaagagga	205080
gccgctgaca	gagcagacag	aaggcaaaga	acaaaagatc	ctccttgggt	agcccgtggg	205140
ctgagggggc	atgggtcagg	tgcttttctt	tcaggctctc	atatgctggg	tgtgggtcca	205200
accagatcca	ctgtagcacg	cacagccaca	gtcagacaca	gtgcatggaa	tgtggaagtg	205260
ctgtgtgtga	gtggaaagtg	gggcttagat	ttagctttca	ggagacagaa	agctccttta	205320
aaagccatac	cttgggctga	ggctgggagt	tggtgtgagt	ggtagagcac	tgtctgtgta	205380
tactcgaggg	ccctgggtgt	cttatctcta	gccccagaag	aagtattaag	aaataaaaagc	205440
aagtgggtgt	tgagatgtga	atggagccag	aactggccgg	aacagtcggg	tggaaagtggg	205500
aagagtgttc	cagacaggga	acagtgtgtg	tgtacctctg	aggctctcat	ggttccatca	205560
gagaggcagg	gaaaggctaa	aatggttttc	ttaagagagt	ccagaagggc	tgggcttggt	205620
ggtgcaggcc	tttaatccca	gcactcagga	ggcagaggca	ggcggatttt	tgagttcgag	205680
gccagcctgg	tctacaaagt	gagttccagg	acagccaggg	ctatacagag	aaacctgtct	205740
tcgaaaaaaa	aaaaaaaaaa	aaagagtgtg	gaagggtgga	agccagggac	aagctctgtac	205800

aagaaggaac	ttgggagcat	tgccgaaagg	atgacctctc	tgcaggtect	gcccaggag	205860
ccagtttctg	gggaccttga	ccatggctag	gtgaatggac	ccaggatggg	atgggtcaggc	205920
ttgctagcag	agccacagcc	gagttggctg	ggtgggggtg	ggtgggggtg	gaaggggtgag	205980
ttatctgatg	agctcaggac	cttttcctgc	cctgcagata	cagcctgcaa	gatgggtccgc	206040
tggctgtctg	ccaagcttgg	ccccacagta	gcctctcgcc	atgtggcccc	gaacctgctg	206100
cgctgctga	catcttgcta	tgttggttaag	gtctgtgggt	agtgtctggag	accaggttcc	206160
ccagccaggc	ttctgccccat	ccttagccct	ctctaggcga	ctccttccct	aacttcccag	206220
cactccctga	gcaggggcctg	ggtctcacc	attaagctgg	gttttcttgg	gtaagtgggg	206280
aagagcccag	tattgaatga	atagaagcca	ccccacagtc	tcagaagggc	ggcttccctc	206340
ctgcctcca	ctggcttctc	aacgctgctg	cccttccctg	gtagggccca	ctcgacagca	206400
gttcacgctc	agcagtgtatg	acacccctcc	actgaatgcc	ggcaacatct	accagaagag	206460
gccagtccta	ggtgacatcg	tgtcggggcc	tgtgtctcagc	tgcctcctcc	acattgccta	206520
cctgtatgga	gaacccgttc	tcacctacca	gtacctgccc	tacatcagct	acctggctcag	206580
tccctgggtc	gtcaaaacccc	ggcttggggg	tgggggcaag	gatccaagga	ccagccccag	206640
gtcttggggg	ttccaggagg	tctgtggggg	gacctgtccc	tccctcatct	attctgtggt	206700
tctaggtagc	cccaggagagc	aactcaaac	ccagccgact	gaacagccgc	aaggaggccg	206760
ggctgctggc	agcggtgaca	ctgacgcaga	aaatcatcgt	atacctctct	gacacgacct	206820
tcattggacat	tctgccccgc	attagccaag	aggtcttgc	gcctgtgctt	ggcttccctca	206880
cctccttcgt	cacagggtag	gcccctgctg	cttgggagag	ccacctggct	gagggggccc	206940
ccaggaagg	ctaggaagct	cagggagaag	cagataccgg	cctgagtc	ggttctgatg	207000
ttgggggtag	tggcacaggt	ctttcattcc	agcaccaga	ggagggcaag	tttctgtgag	207060
tctgagacta	gcctgggtcta	cagagagagc	tccaggctat	ctaaggctcc	atagtaagac	207120
tctgacttaa	gaaaagagtc	gtgggttcatt	ctgggttctg	ggtgtggctt	ggtgatggga	207180
cactttccca	gcatgcagga	ggagctatgc	ttgagttcca	gcccttcaga	aaaacaaaaa	207240
tgggggctgg	aaagaatagc	tcagggttta	agagcactgg	ttgctcttcc	agaggatcca	207300
ggttagattc	ccagctgcca	catggtagct	cataaccatc	cggcagttct	atggaacctg	207360
ccaccctcct	tcgggtctctg	tgggcactgc	aaacatgtgc	acagacatac	atgcaggcag	207420
aaaaaacacc	catacacata	aaattagacc	aaaaaagttc	atgttctctc	ctacctgtag	207480
ctctgactaa	gctacactgc	ttccctgtgc	ctcagtttcc	tcccctggtc	tggactgatc	207540
agccttacat	gcagctcctg	ttatttgaag	ttcctggtaa	atttgtcaag	tccttcaggg	207600
aagggtcggg	aactcttgca	ctttgattct	aggttcccca	gtggggccca	ggccccgact	207660
gtcctatcgc	tgaaaacat	cagtctcatc	gcctctatct	gcttgccat	cgggcaggag	207720
atggtccagc	agcacctgag	tgagccagtg	gccaccttct	tcgaagtctt	ctctcatctg	207780
catgagcttc	ggcagcaggt	aggcaggcag	cttctgggct	gggtgggcca	ggccaggcca	207840
ggccaggcca	gggcagtgga	cccactgaat	ctgtggctct	cctaccgcga	ggatctgcca	207900
ctggatccta	agggctgtac	tgaggggcag	ctgccagagg	cgaccttctc	tgatgggcag	207960
cgacgaccag	tggaccccac	cctgctggaa	gagctgcaga	aggtgttcac	cctggaaatg	208020
gcgtacacaa	tctacgtacc	tttctcctgc	ctgttgggta	ttgcccatca	cgttcctttg	208080
cacagagttg	gtgactacat	ctcttccctg	gggtgggccc	cgatgcttcc	acctccagag	208140
tcagcaatgg	aatcttttta	tttttatatt	gacatggggt	ctcatttagc	ccaggctgac	208200
ctttaactcc	agctccttcc	agcttccacc	gtctcctggt	ggcattgtag	tcattgtggca	208260
ttgctcaggc	ttcttncatg	ttcttatatt	taaatgacct	gtgtgtgtgt	gtgatatctg	208320
tgtgagtgtg	gaggtgacag	aataacagtt	ggggggctcag	cagatgcctt	gcctgatgag	208380
catctctcta	gatccagttt	ttgggtttgt	gggcttttat	gtgtgtgttt	gtttgtctgt	208440
ttttgtagac	aggggtctctc	tgtgtagcct	ggccatcctg	gaactcattc	agtagaccaa	208500
gctggccttg	agctcacaga	gattcacctg	cctctgcctc	ccagtgtctg	gattaaaggc	208560
gtgtaccact	cctgcctggc	tttggttttg	ttaaaccacca	tcctcctgce	tcagcatctg	208620
cctccccctg	gctgggatta	cagggtgtgtg	ctatcacacc	cagctaacag	tggatttaaa	208680
cgtaggaatt	ttaggatcag	agtgaccaga	tttggtccta	gggcccatt	tccacagtga	208740
ttatctatct	tagttaggat	ctctgttgaa	aatcatggtg	gaacatcatt	accaaatgca	208800
acttggggag	gaaaaggttt	attttgtctg	acaactctca	ggtcaccaag	ggaagttagg	208860
gcaggaaactc	gaggcagaag	ctgaagcaaa	agccatggaa	gaactctggc	ttgttccctca	208920
tggcttgctc	agtcctgggtg	accccctccc	caccacctc	cccacaatgg	tttctctgct	208980
tatgcctggc	tgtcctagaa	ctcactctgt	agactaggct	ggcctcaaac	tcaagagatc	209040
cccctgcctc	tgcattctcaa	gtgctaagat	taaaggcggtg	tgccatcacc	cctgccccag	209100
gggtggcact	acccactgta	aattgggtccc	ttcccatatc	agttgttaaa	taagaaaact	209160
cctccatagg	ccaatctggt	gggggaattt	tctcagttga	gggtttctct	tctcaaatga	209220
ctgtagctga	tgccaaattg	ataaaacaaa	tctcaaacca	ccaccaccaa	caacaaataa	209280
accaaacaaa	ccaaacaact	aaccaagaca	gtgacttata	aagagaatct	gaacattttc	209340
cagcaggaaa	ggctcaggag	ctggccattc	aagtctgggg	aacagaatgt	aggggaatat	209400
gatggtctcc	agaagctacc	tgcaaaggaa	tgaacagctt	gctgggtttt	gtggcttccc	209460
ttatgggtag	ggcgctgtac	tgggcttctc	tctgagttag	atgggccacc	ctgtagtgtg	209520
gaatattttg	ctcctacaga	attgtaagtt	ccagaggcca	ggacacatct	ctgttattct	209580
tcattgtgtg	tctgatgcta	gaatggtgcc	tggcatcac	gtgtgtgtct	ctatagagac	209640

agcactcatg	tctacgtatc	gataaaggaa	gctgttttgg	ggggaggaaa	caggcttaca	209700
gacgagaact	taataaccca	gagtagccca	gtcagtagct	tgccttggct	tctgttgttt	209760
ctaagctctg	ggtagatagc	taccttgcca	tcttccctga	tcttagaact	ttccccactc	209820
ccctgtaggt	gacatcatcc	ggaaaatcat	ccccacccat	gagttgggtc	gggagctggc	209880
agggctctat	ctggaaagca	tgagcccagc	ctctcgaaac	ccagccagca	tggaaaccac	209940
catggctagt	gccggccctg	aatgggaccc	tcagagtggg	agctgtctcc	aggacgatgg	210000
ccactcaggg	acctttggga	gtgtcctggg	tggaaatcgc	atccagatcc	ctgactctca	210060
gccccagagt	cctgggccac	tgggctccct	ctctggagtg	ggtagtagcg	gaggcctcag	210120
caacaggaat	gaagacaacg	ccctgaagcg	ggagc'gcct	cggagtgcgc	atgggctgag	210180
cgggaactgg	ctggcgtagt	ggcagtagca	gatcgggtgt	agccagcagg	atgcccaact	210240
ccacttccac	cagatccgcc	tgacagagctt	cccagggcac	acggggggcg	tcaaatgcgt	210300
ggccgccttg	agcagtgaag	acttctttct	gagtggcagc	aaggaccgga	ctgtgcgcct	210360
ctggccgctg	tacaactatg	gggacgggac	caatgagacg	gcttcccgc	tcatctatgc	210420
ccagcaccgc	aaaagcgtct	tctacgtggg	ccagcttgag	gccccgcagt	atgtgggtgag	210480
ctgtgtaggg	gcagtgcacg	tctgggaccc	cttcacaggt	gagcggggcc	aggtgaggcc	210540
tgttcgacgg	ctgctttact	gtgccttagc	caggcctctg	ggaacgggac	ctagtgcgaa	210600
acgtacaatg	gcgtattttg	acggggaaga	ttcagttagg	caggaagaga	agaagagtca	210660
ggacttagaa	tctgtgggac	ccaagtttga	atccactccc	ccaacttacc	agcaatcggc	210720
tcagttagctg	caggcgtctg	ccttctacct	gtaagaacca	aaaattttaga	agattccacg	210780
agtagggctt	tggcttcttg	tacgacgtca	cctgtcgtcg	ttgtaaagag	aagtatcgag	210840
tggaggaggg	tcagggcaga	cggaggtcgc	agcatgctat	agcagagtag	tgtaagagag	210900
cagactgttc	tggggctgga	cccttgactt	cactgtggaa	gcagcaagat	gagaaagccc	210960
tgagattgtg	ttttctgagg	gtcactgggg	aatgggatgc	aggtgtgggg	tgagttggag	211020
tttgaagtag	ccagggctct	ttgatagcca	ctaagtcccc	agatgtgtcc	tttttcagga	211080
aagacccttc	gcacagtgga	tccttcagac	agccgggtgc	ccctgacggc	tgtggctgtc	211140
atgcctgccc	cacacaccag	catcaccatg	gccagctccg	actccactct	gcgctttgtg	211200
gactgcagga	agccaggctt	gcaggtcagg	aggggtgcag	ttcctgggct	actgggggtc	211260
tctaggtacc	agtcaggaaa	gacactcagg	ggactccacc	aggaacgctg	cagtgcaggg	211320
cagccctgtg	tgggtggggc	gctggcacgg	atggggcttt	tctcttcccg	ggatggagtg	211380
ggagggtcag	gcctactggt	ttcgtggggc	tgaatggggg	gagctgcagt	aggggtgggtg	211440
gcagtgatgg	atggcgacgg	gcacttgaac	acaatctcct	cctatagcat	gagttccgac	211500
tgggtggagg	gctgaaccct	gggcttgttc	gctcgttggc	cgtcagcccc	agtggccgga	211560
gtgttgtggc	tggcttctcc	tcgggcttca	tgggtgctct	agatacccgc	acgggcctgg	211620
ttctacgagg	ctggccagcc	catgaagggg	acattctaca	gatcaagggtg	actgactgcc	211680
tgaggtccta	tcctttcatt	tctacttagg	gcctgggtctg	ggagaggaca	ggtttatgct	211740
ggtgtccctt	ataactactc	ggggacattc	agtgggggtg	gaaaatggcc	ctcgtaggcc	211800
agctcaggaa	ccagctgcac	aggaggcagg	ctaggggcag	gaatcagggc	tagaactgac	211860
cctgatgtct	cacagcgatg	ttctaattgag	taacccttgt	ccatatattgt	cttgccttga	211920
ggatcagggg	tcacgcccctg	tccgtgaccc	agttcaggtt	aaataaagcc	aggaggctgt	211980
ttactgcctg	gagaccactg	agcagagctc	atgccccctg	ctgggctgtc	ctgatggggg	212040
gcaggaacag	gcgcaggcct	gcgcacgtgt	ttcctgcctc	ctatatccaa	tcatagacct	212100
cagagctcag	caggttctgg	gaggggagaa	atagggctgc	ttgtgggagg	atttctccct	212160
gcagtgggaa	ctctcctccc	ccgccttcca	gatggaggtg	aaagacaggc	actgttgctt	212220
acagggaagg	caggctgccc	ccagctctat	ccaggacccc	aggggaccct	gggtctcagt	212280
gtctctaaat	cccaacattc	taagaaagtg	tcaggatggc	tctggggctc	tcctgggtgt	212340
tagtcccagc	tctcggagtc	ttctctgagc	accagttctt	ttcctatggg	aagtaaggac	212400
atgccagggtg	ttctttgaga	ggacctgagt	ttggttctca	gcactgtcta	gctctggctc	212460
caggggggttc	aacacccttt	atggcttctg	tggacacata	ttcttatgtg	gcctgcacgc	212520
acacacacga	acacaaataa	aaataaatgt	taaaagaaga	cagcggcacc	ttgtacctca	212580
catgttagta	cgattggatg	tggcagtgcc	tcacagaatc	tgtgggactt	tattttattta	212640
tttatttttt	ggtcttaaaa	ttttaaaaga	ttggtttagg	agtgggtgga	cacaccatta	212700
atcgcaacac	tcaggagcag	aggcaggtgg	atctctatgg	gtttgaggcc	agcctggctc	212760
acagagcaag	tttcagggca	gccaagggtta	cacagagaaa	ctctttctca	aaaaataaaa	212820
acaaaatatt	taaaagattt	acttacttct	tatttgagct	aggatctccc	tatttagccct	212880
ggctgtcctg	gaactcactg	tatagaccag	gctggcacct	taaactcacg	aagatcctcc	212940
tgcttctgcc	tcctaagtgc	tgagattaa	gtagtgttat	accatgcccc	actattttct	213000
ttatagattt	ggtgttttgc	ctgcttgtgt	atatatgcac	taccttcacg	cagtgctaat	213060
aggggtcaga	ggcgagtatc	agctcttctc	ggaactagag	ttatggaagg	ttgggagtca	213120
ccatgctggg	actgggtcat	ttgcaagagt	tacaagtact	tctgagccat	ctccagcccc	213180
ctagagtttt	tttccccctt	ggctgtcctg	aagtagaatc	tgttcttgtt	ttgttttgtt	213240
tttcgagaca	gggtttctct	acataggcct	ggctgtcctg	gaactcactc	tgtagaccag	213300
gctggcctcg	aactcagaaa	tccgcctgct	tctggctctc	agaatactgg	aattaaaggt	213360
gtgcccacc	acgcctggct	cagaatctgt	tttttaaatga	gagtaatagt	tacaggtttt	213420
ttgttttgtt	tttttctttt	ctttttttgt	ttttgttttt	tggctcattt	gtttttattt	213480

```

ttttgagaca ggtctcactc tgaaccctta ggtggcctgg agcttgctat gtagaacaca 213540
ctgacttttaa acttggttttc tgagtgtctgg atttatgggc ttgtgctatt ttgccagacc 213600
tctgatgggtt gttaataaca atattatttta gcttttcttt tggagatagg ctctcactgt 213660
atatcaccca gacagtggct ggtctggaaa tcaactgtga ggccaggctg accttgaatt 213720
cacagagatc tgcctcccgga gttcagaaat taaaagcact ctgggatgggt ttggagttt 213780
ggtgagtacc caagcctcca ttgatgtctat ctgtccctcc cgctctctgc aggctgtaga 213840
gggcagcgtg ctcatcagct cctcttccga ccattccttg actgtttgga aggagctgga 213900
acagaagccc acgcaccact acaagtcagc gtccgaccca atccacacct ttgacctgta 213960
cggcagcgag gtggtcaccg gcaactgtagc caacaagatt ggtgtctgtt ccctgcttga 214020
gccaccctct caggccacca caaagctcag ttccgagaac ttccgtggca cgctcactag 214080
tctggctttg ctgcccacga aacgccacct cctgctgggc tcggacaatg gcatcatccg 214140
cctcctggca tagggccagc caggagtgg ctgagggcag ggcgagatga catctctcag 214200
ggcccgctcc tcattcttga tctcgaagcc gattcttcta ggcaagcccc aggctctggc 214260
taccacatg gcctgctgtc tgggattgca cagctcctga atctccaaag ccttgaagtg 214320
gcttcatgaa actcgggaga tactgttcct aaccagcaag aattggggca aggaaagcac 214380
tgtgatcccc attgctcccc agttctgcct tctggattca catggggaca gggcagctcc 214440
aggaaatgaa aggagtggg cctttgctca gccagcttcc tctagccacg ctctccttag 214500
ctctgtttct cccttgggta ggaaactgct cctgtctagg gttctgatgg tactgggact 214560
ccaggctcag gagggctggc caggacctac gactttcagg gcttgggtctg gggttttagc 214620
attcattcag ccaggctctc agtatgggac cagaaaaaag gggatgtgag aacagggtcta 214680
gggaagggggt tatatggggc cagctggctc aggaatgaat ccatgccttg ccttgggtacc 214740
cctaaccaca gcgtttgtgc cttcagccgg ggaggcagcc cttgggacca gcatccctag 214800
ggacaggagg cagcgggaat catctctgta tctcgggttc tgcccagggg atgggcagac 214860
tctgccatct cttgagtgtt cgtttgagga agcctgagat gtggccccctg ctgccttctc 214920
actagtgtca gtctatgtaa ataaggtcaa taaattcttt ggaagagcca cggagctgag 214980
tgaggctgtg ttgtgttttg ctttgcttag gctgggctca ggcagctctg cctcagcctc 215040
ccaaggagct ggggaactgg tatatgtcac tgtatatgtc actgtgcctg gcttatggct 215100
tggcttggct ttttttcaga tgggtctcaag tgccctcagg ttggccttgat cttgggatga 215160
ccttctgctg tgaaacagag tagtgggctt ataggcatga cccaccagggt ccaattttta 215220
ttttttaaaag gcattgattt ttatacgtgt atgggtgttt tgcccacttg tacatatgca 215280
caccatactt gtgtctgggc cctgcggagg tcagaagagg gcatcgggat cacctggaaac 215340
cgaagttaat gaatggttat gagccacatc tcgatgctga agattgaacc tggatccctt 215400
gcaagagcag ccagtgttct taccactga gccatctcta agccccacac ccagcttctt 215460
ttgatacaag gtctggtagc tcaaacttga tatgcagccg aggaggttga cctgggtattc 215520
cctacctaacc ctcttctctc taccttccaa gtgctgatat tatacatagg catggatagt 215580
catgccacc agtttgctt gatggcacca gagtccaggaa agtccaaacc ggtagttgc 215640
aaacacagca agagggtaga ggcagccatt gtccctctggc tgccttggat acagagcttc 215700
tgggttgggt ggccttgggt cagttttccg aatgggttcac ccttggggaa agggaaact 215760
gctgaagagg tgggaccctg ggagggcccg cctccagctg ggtctctcca gccctcgct 215820
tggaacctag gctggagggga gccaaaccagg atcctggact tgctacagtt aggtgaacag 215880
gctcctgcag cctccccttc ccttgggtag ctgtggtggt ggtggtggtg gtggtggtg 215940
tggtggtggt ggtggtggtg gtgggggggg gggngnngnt 215980

```

<210> 17

<211> 473

<212> PRT

<213> Mus sp.

<400> 17

```

Met Lys Arg Ala Ser Ser Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
  1             5             10             15

```

```

Trp Leu Gln Ala Trp Arg Val Ala Thr Pro Cys Pro Gly Ala Cys Val
          20             25             30

```

```

Cys Tyr Asn Glu Pro Lys Val Thr Thr Ser Cys Pro Gln Gln Gly Leu
          35             40             45

```

```

Gln Ala Val Pro Thr Gly Ile Pro Ala Ser Ser Gln Arg Ile Phe Leu
          50             55             60

```

```

His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Gln Ser Cys
          65             70             75             80

```

Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Ala Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Thr Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu His Val Val Asp Pro Thr Thr Phe His Gly
 115 120 125
 Leu Gly His Leu His Thr Leu His Leu Asp Arg Cys Gly Leu Arg Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Asn Leu Gln Ala Leu Pro Asp Asn Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Pro Ser
 180 185 190
 Val Pro Glu His Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu
 195 200 205
 Leu His Gln Asn His Val Ala Arg Val His Pro His Ala Phe Arg Asp
 210 215 220
 Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Met
 225 230 235 240
 Leu Pro Ala Glu Val Leu Met Pro Leu Arg Ser Leu Gln Tyr Leu Arg
 245 250 255
 Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp
 260 265 270
 Ala Trp Leu Gln Lys Phe Arg Gly Ser Ser Ser Glu Val Pro Cys Asn
 275 280 285
 Leu Pro Gln Arg Leu Ala Asp Arg Asp Leu Lys Arg Leu Ala Ala Ser
 290 295 300
 Asp Leu Glu Gly Cys Ala Val Ala Ser Gly Pro Phe Arg Pro Ile Gln
 305 310 315 320
 Thr Ser Gln Leu Thr Asp Glu Glu Leu Leu Ser Leu Pro Lys Cys Cys
 325 330 335
 Gln Pro Asp Ala Ala Asp Lys Ala Ser Val Leu Glu Pro Gly Arg Pro
 340 345 350
 Ala Ser Ala Gly Asn Ala Leu Lys Gly Arg Val Pro Pro Gly Asp Thr
 355 360 365
 Pro Pro Gly Asn Gly Ser Gly Pro Arg His Ile Asn Asp Ser Pro Phe
 370 375 380
 Gly Thr Leu Pro Ser Ser Ala Glu Pro Pro Leu Thr Ala Leu Arg Pro
 385 390 395 400
 Gly Gly Ser Glu Pro Pro Gly Leu Pro Thr Thr Gly Pro Arg Arg Arg
 405 410 415
 Pro Gly Cys Ser Arg Lys Asn Arg Thr Arg Ser His Cys Arg Leu Gly

420 425 430
Gln Ala Gly Ser Gly Ala Ser Gly Thr Gly Asp Ala Glu Gly Ser Gly
435 440 445
Ala Leu Pro Ala Leu Ala Cys Ser Leu Ala Pro Leu Gly Leu Ala Leu
450 455 460
Val Leu Trp Thr Val Leu Gly Pro Cys
465 470

INTERNATIONAL SEARCH REPORT

Inte. Application No

PCT/US 01/31488

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/705 C07K16/28 C12N15/62 G01N33/68
A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBL, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL NUCLEIC ACIDS 'Online! 21 January 2000 (2000-01-21) "M. musculus chromosome 11 clone RP23-113D4" retrieved from EMBL Database accession no. AC021768 XP002217414 compare translation product of complementary strand of nt 153209-154529 with aa 21-461 of seq. 4 and aa 1-392 of seq. 14</p> <p style="text-align: center;">--- -/--</p>	1,2,5-30

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

18 October 2002

Date of mailing of the international search report

08/11/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Galli, I

INTERNATIONAL SEARCH REPORT

Internat Application No

PCT/US 01/31488

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL NUCLEIC ACIDS 'Online! 15 November 1999 (1999-11-15) "Homo sapiens clone RP11-10E2" retrieved from EMBL Database accession no. AC013606 XP002217419 compare aa 168-420 of seq. 2 with the translation product of nt 138578-139331.</p>	1-4,6-30
X	<p>DATABASE EMBL PROTEINS 'Online! 1 May 2000 (2000-05-01) CORBY N.: "New leucine-rich protein" retrieved from EMBL Database accession no. Q9UGS3 XP002217415 compare aa 352-399 with seq. 11,19</p>	1,11,15
A	<p>DATABASE EMBL PROTEINS 'Online! 1 October 2000 (2000-10-01) OSADA N. ET AL.: "Unnamed protein product" retrieved from EMBL Database accession no. Q9N0E3 XP002217416 compare with sequences 2,4,5,11,17,18,19</p>	1-30
A	<p>WO 00 31235 A (CHEN MAIO S ;SCHWAB MARTIN E (CH)) 2 June 2000 (2000-06-02) the whole document</p>	1-30
P,A	<p>WO 01 51520 A (UNIV YALE ;STRITTMATTER STEPHEN M (US)) 19 July 2001 (2001-07-19) NgR1: compare seq. 1,3 with seq. 5,17 of present application, respectively</p>	1-30

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims 1,2,6,7,9-13,15,16,18-30 relate to an extremely large number of possible compounds. It is not possible to determine all the possibilities, as any given X can be any amino acid or even a gap. In fact, the claims contain so many options, variables, possible permutations and provisos that a lack of clarity (and/or conciseness) within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible. Consequently, the search has been carried out for those parts of the application which do appear to be clear (and/or concise), namely seq. ID 2, 4 and 14.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/31488

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 27, 29 and, as far as methods in vivo are concerned, claims 26 and 28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/31488

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0031235	A	02-06-2000	AU 1469200 A	13-06-2000
			CN 1354755 T	19-06-2002
			CZ 20011608 A3	17-10-2001
			EP 1124846 A2	22-08-2001
			NO 20012223 A	02-07-2001
			SK 6222001 A3	03-12-2001
			WO 0031235 A2	02-06-2000
WO 0151520	A	19-07-2001	AU 2940101 A	24-07-2001
			EP 1248803 A2	16-10-2002
			WO 0151520 A2	19-07-2001
			US 2002012965 A1	31-01-2002
			US 2002077295 A1	20-06-2002

Form PCT/ISA/210 (patent family annex) (July 1992)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record.**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.